



HAL
open science

Concepts : Learning and Generalization : a developmental approach

Eleanor Stansbury

► **To cite this version:**

Eleanor Stansbury. Concepts : Learning and Generalization : a developmental approach. Psychology. Université Bourgogne Franche-Comté, 2023. English. NNT : 2023UBFCH037 . tel-04651198

HAL Id: tel-04651198

<https://theses.hal.science/tel-04651198v1>

Submitted on 17 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**THÈSE DE DOCTORAT DE L'ÉTABLISSEMENT
UNIVERSITÉ DE BOURGOGNE**

École doctorale n°554 - École doctorale Environnement Santé

Doctorat de Psychologie du Développement

**Concepts: Learning and Generalization.
A developmental approach.**

Concepts : Apprentissage et Généralisation. Une Approche Développementale.

Par

STANSBURY Eleanor

Thèse présentée et soutenue à Dijon le 12 décembre 2023.

Composition du Jury :

Pr. Elisabeth Demont	Professeur Université de Strasbourg	Examinatrice – Présidente du Jury
Pr. Jérôme Clerc	Professeur, Université Grenoble Alpes	Rapporteur
Pr. Nathalie Blanc	Professeur, Université Paul Valéry Montpellier 3	Rapporteur
Pr. Jean-Pierre Thibaut	Professeur, Université de Bourgogne	Directeur de thèse
Dr. Arnaud Witt	MCU, Université de Bourgogne	Co-directeur de thèse

To James

*« La vie, ce n'est pas attendre que l'orage passe,
c'est apprendre à danser sous la pluie »
Sénèque*

Titre : Concepts : Apprentissage et Généralisation. Une approche développementale.

Mots Clés : Comparaison, Généralisation, Eye-tracking, Concepts, Stratégies, Nom, Nom relationnel, Catégorisation, Taxonomie, Perception, Théorie de la détection du signal, Fonctions exécutives, choix libre, choix forcé.

Résumé : Une capacité que nous considérons comme acquise – mais qui n’en est pas moins remarquable – est notre capacité à nommer correctement des objets que nous rencontrons pour la première fois, dès lors qu’ils appartiennent à une catégorie d’objet que nous connaissons. Cette capacité qui repose sur nos concepts et notre habilité à généraliser les noms entre objets de même catégorie est indispensable tant pour se repérer que pour communiquer. À ce jour, les recherches ont très largement mis en évidence que l’apprentissage de mots à partir de plusieurs exemples d’objets à comparer – dit apprentissage en situation de comparaison – favorise la généralisation correcte de nouveaux mots par rapport à des situations d’apprentissages avec un exemple unique. Dans le présent travail, après avoir revu les études de généralisation de nouveaux mots en situation de comparaison et leurs propositions quant aux mécanismes qui sous-tendent la généralisation de nouveaux mots, nous posons la question : *comment* généralise-t-on les nouveaux mots ? Nous nous interrogeons sur les stratégies de généralisation utilisées, leur organisation dans le temps et les mécanismes qui les sous-tendent. Nous abordons cette question sous deux angles différents et complémentaires. Dans un premier temps, nous explorons avec une approche développementale les stratégies de généralisation de nouveaux noms à l’aide du suivi du mouvement des yeux (eye-tracking) pour révéler le profil des stratégies de généralisation de nouveaux noms. Puis nous considérons la généralisation de nouveaux noms dans un paradigme nouveau, à choix libre plutôt que forcé, pour modifier les contraintes qui s’exercent sur la tâche et étudier toutes l’étendue des généralisations considérées possibles par les enfants. Ces études révèlent pour la première fois des profils systématiques de stratégie de généralisation dont l’efficacité de la mise en œuvre varie au cours du développement. Elles proposent aussi une interprétation de la généralisation de nom en distinguant la sensibilité des sujets aux exemples de même catégorie et, ce qui est nouveau, leur inclinaison à généraliser selon une règle de catégorisation donnée.

Title: Concepts: Learning and Generalization. A developmental approach.

Keywords: Comparison, Eye-Tracking, Concepts, Strategies, Nouns, Relational nouns, Categorization, Taxonomy, Perception, Signal detection theory, Executive functions, free choice, Forced choice.

Abstract: A remarkable ability that one may take for granted is the ability to name objects encountered for the first time when these are members of known categories. This is possible thanks to our conceptual system and ability to generalize names between same-category members. It is priceless for organizing the world around us and communicating. Research has shown that learning words in a comparison setting – with multiple items associated with the novel word rather than one – favors correct generalization. In the following work, we first review the literature on novel word generalization and their suggestions about the mechanisms that may underlie it. Second, we investigate *how* one generalizes novel words. We question which strategies are used to generalize novel words, what are their temporal dynamics, and which processes may underlie them. We address this question in two different and complementary ways. We investigate novel noun generalization strategies with eye-tracking data to reveal the strategies' profiles and their developmental course. Then, we investigate novel noun generalization performances in free-choice rather than forced-choice design to change the task's constraints and consider a broader scope of children's generalizations. These studies reveal, for the first time, clear novel noun generalization strategy patterns. The patterns are stable across development: it is how the strategy is implemented that that develops. The studies also suggest that novel noun generalization may be considered with two components: children's sensitivity to same-category items and willingness to follow a given categorization rule.

REMERCIEMENTS

Tout d'abord je souhaite remercier l'ensemble des membres du jury de cette thèse pour leurs temps et leur expertise scientifique. Merci à Nathalie Blanc et Jérôme Clerc de m'avoir fait l'honneur d'accepter d'être rapporteurs et pour leurs conseils éclairés. Merci aussi à Elisabeth Demont pour sa participation en tant qu'examineur.

Je tiens à remercier mon directeur de thèse Jean-Pierre Thibaut, pour son accompagnement et son exigence durant ces années de thèse. Je tiens aussi à remercier mon codirecteur de thèse, Arnaud Witt, pour ses conseils et sa clairvoyance. Je leur suis grandement reconnaissante pour la formation qu'ils m'ont transmise et l'expertise dont ils m'ont fait bénéficier.

J'adresse un grand merci à Micah Goldwater et Jean-Michel Boucheix pour avoir accepté de faire partie de mon comité de suivi de thèse et pour avoir apporté leur avis et leur recul tout au long du travail. Nos échanges ont toujours été «*great food for thought*» et des échéances importantes pour envisager la poursuites des travaux.

Je souhaite remercier l'ensemble de l'équipe de l'École Doctorale Environnement et Santé, pour la qualité de son encadrement, et de ses formations. Je remercie aussi le Ministère de l'Enseignement Supérieur et de la Recherche et l'École Doctorale pour la bourse dont j'ai bénéficiée tout au long de mon parcours.

Un grand merci à l'ensemble des écoles qui m'ont accueillie durant ces quatre années, ainsi qu'aux enfants et à leurs parents qui ont accepté de participer à ce travail. Sans eux, rien n'aurait été possible.

Je remercie chaleureusement tous les membres du LEAD qui m'ont accueillie durant ces quatre années, et qui m'ont offert un cadre de travail stimulant. Je remercie Bénédicte Poulain-Charronnât la directrice du LEAD. Je remercie l'ensemble des doctorants et jeunes chercheurs pour les discussions, les réunions jeunes chercheurs, et les sorties qui ont animées notre temps au laboratoire. Merci à Yannick, Damien, Iva, Claudia, pour nos collaborations et pour avoir montré la voie. Merci à Florian, Laetitia, Julie et Williams pour vos cafés joyeux et votre bonne humeur. Un grand merci aussi à Julie Ferreira pour nos échanges. Merci enfin à Sandrine et Corine pour votre travail qui fait vivre le labo.

Et enfin je souhaite remercier mes proches qui m'ont soutenue et aidée durant ce travail. Vos conseils pour entreprendre ce travail et vos encouragements tout au long m'ont été indispensables.

INDEX

REMERCIEMENTS	4
INDEX.....	6
SCIENTIFIQUES CONTRIBUTIONS ASSOCIATED TO THE THESIS	8
GENERAL INTRODUCTION.....	10
SECTION 1: A THEORETICAL REVIEW	30
Apprentissage et généralisation du lexique par comparaison : les mécanismes de l'expansion conceptuelle.	32
SECTION 2: INVESTIGATING WORD GENERALIZATION STRATEGIES.....	70
Introduction.....	72
Chapter 1 - How children generalize novel nouns: an eye-tracking analysis of their generalization strategies.	78
Chapter 2 - Relational noun generalization strategies: A developmental eye-tracking analysis.	136
Chapter 3 – Investigating cognitive mechanisms that underlie novel word generalization.	168
SECTION 3: DIFFERENT INSIGHTS INTO GENERALIZATION WITH A DIFFERENT PARADIGM	190
Introduction.....	192
Chapter 1 - Generalization of novel object names in comparison contexts in a yes-no paradigm by young children. When the rate of stimulus presentation matters.	202
Chapter 2 - How stimuli availability effects novel noun generalization in a free-choice design.	214
Chapter 3 - Novel noun generalization in a free-choice design: investigating generalization constraints.	224
GENERAL DISCUSSION	234

SCIENTIFIQUES CONTRIBUTIONS ASSOCIATED TO THE THESIS

Section 1

Lagarrigue Y., Stansbury E., Thibaut J.-P. (submitted). Apprentissage et généralisation du lexique par comparaison : les mécanismes de l'expansion conceptuelle.

Section 2.

Chapter 1. Stansbury E., Witt A., Thibaut J.-P. (submitted) How children generalize novel nouns: An eye-tracking analysis of their generalization strategies. *PlosOne*

Chapter 2. Adapted from: Stansbury E., Witt A., Thibaut J.-P. (2021) Children's Generalization of Novel Relational Nouns in Comparison Contexts: An Eye Tracking Analysis. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43th Annual Meeting of the Cognitive Science Society* (pp. 2808 - 2814).

Section 3.

Chapter 1. Stansbury E., Witt A., Thibaut J.-P (2020) Generalization of novel object names in comparison contexts in a yes-no paradigm by young children. When the rate of stimulus presentation matters. In S. Denison, M. Mack, Y. Xu, & B. Amstrong (Eds.), *Proceedings of the 42th Annual Meeting of the Cognitive Science Society (Eds.)*, (pp. 2288 - 2294).

Chapter 2. Stansbury E., Witt A., Thibaut J.-P (2022) How stimuli availability effects novel noun generalization in a free-choice design. In J.Culbertson, A.Perfors, H.Rabagliati & V.Ramenzoni (Eds.), *Proceedings of the 44th Annual Meeting of the Cognitive Science Society*, (pp. 2789, 2795).

Chapter 3. Stansbury E., Witt A., Thibaut J.-P (2023) Novel noun generalization in a free-choice design: investigating generalization constraints. In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society*. (pp. 3345, 3341).

GENERAL INTRODUCTION

General introduction

While many theories exist about learning words, they all admit that words must be learnt and question how this is done. They differ on the amount of prior knowledge needed to learn words for example, but all consider that children must learn item-word associations to acquire language.

Learning item-word associations is understanding which items are named by which words. Children must complete this object to symbol (i.e. word) mapping process and remember the association. This works for nouns that are associated to concrete, identifiable items. However, children must go beyond individual memories of object-symbol associations to learn the full extent of language. To learn other types of words (i.e., verbs, relational nouns, adjectives etc.), that are not associated to such a definable item in our environment, children must learn words' meanings (Bloom, 2002).

A word's meaning may have numerous structures depending on the word itself and the individual's understanding. A noun's meaning is a representation of the object the noun refers to, an adjective's meaning is a representation of the characteristic the noun refers to, a verb's meaning is a representation of the action the verb describes etc. In all cases however, when the meaning is stored in memory, it is stored in the shape of a mental representation, a concept.

The difficulty for children that are learning word meanings is therefore to build the right concept for the right word. This word-to-concept mapping process is frequently biased in young word learners. For children's word-to-concept mapping to be correct, they must map words to item's conceptual meaning rather than other item characteristics such as color, shape, function, belonging etc. Children, however, tend to map words to the item's characteristic that they find the most salient until they learn to map on a taxonomic basis. This tendency leads early word learners to extend words not to taxonomically related items but to items with the same shape (shape-bias, Landau et al., 1988b; Smith et al., 1996), to items often found in the same context (thematically related extensions, Gentner & Namy, 2004; Gentner & Rattermann, 1991), to items of the same basic level category only (basic level bias, Waxman & Hatch, 1992; Waxman & Kosowski, 1990). How children extend the words they have learnt in each of these cases reveals the type of representation they have built for the word's meaning. Here, in the case of biases, the representations are based on dimensions that are not conceptual dimensions and lead children to misunderstand the words and generalize them incorrectly. It is therefore crucial that children build conceptually based concepts, to grasp word meanings correctly and use words efficiently when communicating.

General introduction

An important issue in the domain of word learning, and more broadly conceptual development, is thus understanding how children build conceptually based concepts associated to novel words. This issue may be considered from the learning situations. The question is then: which word learning situations help children identify items' conceptual dimensions, to build conceptually based concepts and generalize words correctly (i.e. to conceptually related items)? This question has been extensively studied. Inferences about how children learn words have been made from analyzing variations in their word generalizations as a function of the learning situation's parameters. Here, we built on this body of research.

The question we pursued was: *how* do children come to a conceptual generalization decision? This led us to investigate generalization strategies, their profiles and their temporal dynamics, and the processes that may underlie noun generalization. We addressed this issue in three sections. In the first section, we introduce a theoretical review of word learning and generalization in comparison settings. A large body of research has shown that comparison settings, in which children learn a novel word from two exemplars rather than a single exemplar, favor conceptual word generalization. We will review this research in the case of nouns, relational nouns, verbs and adjectives. We will summarize the word generalization processes that have been inferred from these empirical studies and that are our starting point into the investigation of how children come to a conceptual generalization decision. In the second section, we present two real time investigations of children's mapping processes in the case of noun and relational noun generalization. We tested children in noun and relation noun generalization tasks while recording their eye movements to get a better understanding of the temporal dynamics of their word generalization strategies. We also aimed at investigating the cognitive processes that underlie children's generalization strategies. This sections last chapter summarizes this preliminary investigation. In the third and last section, we present a series of novel noun generalization studies in free choice designs. Novel noun generalization has most often been studied in forced choice designs, where children must choose the item they consider best fits the newly learnt word. However, these designs constrain generalization choices, and it is interesting to investigate children's generalization decisions when this constraint is lifted. In forced choice designs in which children must choose a single answer, they may exclude multiple items that they might have extended the novel word to as well as the best fit item, or they may generalize to an item that they might not have chosen otherwise. In the studies presented in this section we consider generalization decisions in traditional way, by inferences from what is observed of children's choices, but in a more ecological and informative setting.

Before coming to these sections however, it is necessary to consider what concepts are and how they develop. We will review these questions in the following paragraphs.

Concepts: importance and definition

Concepts are a fundamental cornerstone of human intelligence. They are essential for anything we want to do, plan, say, or remember. Concepts are the mental representations that we have of the world that surrounds us. They tell what things are and what properties they have (Murphy, 2002), enabling us to encounter novel objects, people, events, while still understanding the world around us. For example, when we enter an unknown conference room, we know which items inside the room are chairs, tables and electronic devices even though we have never seen these specific items of furniture and devices before. We also know that the chairs are for sitting, that we can sit in front of the tables and that electronic devices need to be turned on to be used. All these simple actions would be impossible without previously built and stored concepts of the categories “chair”, “table” or “electronic device”. Indeed, the main roles of categories are to help classify the environment and serve other cognitive functions especially inference and communication (Anderson, 1990).

Even though, the distinction between category and concept is not always clear in psychological research, because these terms cover very similar meanings, the term category usually refers to a collection of *items* that are treated as equivalent, whereas a concept is the *mental representation* of a category or of the items the category encloses (*APA Dictionary of Psychology*, s. d.; A. B. Markman & Rein, 2013; Murphy, 2002). We shall use the words category and concept in these ways.

Concepts are building blocks of human cognitive activity and are consequently a key research topic. In the present work we will study novel noun learning and generalization as an example of concept learning and use. Our aim is to get a better understanding of how mental representations are built and how they are later used to make inferences about the world - and pursue learning.

In the following paragraphs, the framework in which our work can be considered is presented, arguing for the dynamic nature of concepts that change as a function of an individual's development and experience, and of concept learning and generalization contexts.

Building concepts

Models of concepts

Many models of concepts have emerged in the psychology literature in the past decades such as the classical view (Hull, 1920; Mervis, 1980), the prototype view (Mervis & Rosch, 1981; Reed, 1972; Rips, 1975; Rosch et al., 1976; Rosch, 1973), the exemplar view (Brooks, 1987; Medin & Schaffer, 1978), the knowledge view (Keil, 1992; Murphy & Medin, 1985), and more recently the rational approach (Anderson, 1990), decision bound models (Ashby & Maddox, 1993) and neural network and connectionist models (Knapp & Anderson, 1984; Krushke, 1992). Despite their many differences, that we will not develop in detail here, all these models agree that concepts are reduced, more abstract, schematized representation, used for future inferences (Rips et al., 2012; Sloutsky, 2010). Over time, the models' complexity and how dynamic concepts are considered has progressively evolved.

The classical view defined concepts as definitions of the necessary and sufficient characteristics for category membership (Hull, 1920; Mervis, 1980). This suggestion implied that categories have rigid category member boundaries: items were either a category member or not. The rigidity of this view, that couldn't account for empirically revealed typicality effects or "border line" category members (Mervis & Rosch, 1981; Murphy & Brownell, 1985; Rosch et al., 1976b; Smith & Kemler, 1978), was the cause of the view's downfall.

Two new views emerged from this downfall of the classical view, the prototype view (Mervis & Rosch, 1981; Reed, 1972; Rips, 1975; Rosch et al., 1976; Rosch, 1973), and its immediate counter proposition: the exemplar view (Brooks, 1987; Medin & Schaffer, 1978). The groundbreaking conception common to both these views is that concepts are mental representations built of category defining *features* and that category membership is decided upon after a comparison between an item's features and the representation's features. Representation features were weighted (by importance for category membership in the prototype view, as a consequence of frequency of experience in the exemplar view) and an item is judged a category member if the similarity between the item and the representation reaches a certain threshold on a continuum. Consequently, both views account for typicality and similarity effects, and opened the way to more dynamic ways of considering concepts.

The views' difference lies in the complexity and the rigidity of the mental representation. The more ridged view, the prototype view, postulates that individuals build a summary representation of the category items, that encompasses an entire concept, and that category membership is assessed by the

General introduction

similarity between an item and the prototype following a feature additive rule (Tversky, 1977). The exemplar view takes a broader approach and postulates that the representation is built from all features experienced with items from the category and stored in memory. In this view, one way of considering categorization, has been by considering that category membership is assessed by the similarity between the item and the stored memories. The similarity assessment would then follow a multiplicative rule that considers both feature similarity and the importance of the similar features for category membership (Medin & Schaffer, 1978; Murphy, 2002). Thus, what the exemplar view adds to the prototype view, is the broadness of features in the representation and the suggestion of an interplay between the mental representation and the individual's experience and memory systems.

The slightly later view, the knowledge view (Keil, 1992; Murphy & Medin, 1985) also grants concepts with a dynamic nature, in interplay with individuals existing representations. The knowledge view considers that concepts are part of general knowledge. This has two implications, first that concepts are not learnt in isolation from what one already knows and are thus influenced by what is already known when they are learnt, and second that new concepts influence one's general knowledge. This view adds an important point to our understanding of conceptual development by highlighting how learning concepts is influenced by one's existing knowledge.

We can note here that the different theories sparsely discuss the nature of the features upon which representations are built and don't agree on this point. We will discuss later that these can be perceptual features or lexicalized - more abstract - features. The proportions of perceptual and lexicalized features in concepts, and their relations, lead to different types of concepts (Sloutsky, 2010).

To summarize the points of agreement of these numerous theories of concepts, we can say that concepts are mental representations with weighted features, to which items are compared to decide on their category membership. This leads to our next question, how are these mental representations built?

Building a mental representation

Concepts are built as information from the environment is selected, reduced and stored in memory. A key to concept building is therefore the ability to focus on relevant aspects of the information (i.e., information that defines the category) and ignore others (Sloutsky, 2010), so that the information integrated into the concept is the most pertinent for defining the category.

Many different types of information can define the many category types observed in adults. Markman & Rein's (2013) description of categories illustrates the diversity of categories and

General introduction

categorization rules. They describe five different types of categories: feature based, goal derived, relational, role governed, thematic. Feature based categories are categories represented by common features; relational and role governed categories are represented by common relations among items; thematic and goal derived categories are organized around items use or function. Items can thus be grouped in a category defined by any of their communalities from perceptually grounded features to their most abstract properties.

In the present work we will focus mainly on feature-based concepts (i.e., perceptual categories, taxonomic categories) and to a lesser extent to relational concepts. We focus mainly on one type of feature-based category: taxonomic categories, because they are prevalent in communication, and their acquisition and use are a key focus point for young children's conceptual and language development.

Building taxonomic categories

A taxonomic category is a feature-based category that includes all conceptually equivalent items. The defining features of taxonomic categories are therefore conceptual linguistic information (Murphy, 2016), even though same taxonomic category items in the human concept structure share multiple perceptual features and thematic relations.

This is the source of difficulty when learning taxonomic categories: to build taxonomic categories (i.e., notice taxonomic features to be abstracted and stored in the concept) one must focus on conceptual information rather than more salient perceptual information or more frequent context or function related thematic relations. This requirement has a few implications for building taxonomic categories.

First, selective attention mechanisms must be efficiently directed to allow focusing on conceptual information. Second, participants must be able to process linguistic information that are part of taxonomic concepts (e.g., names, lexicalized features, learnt category hierarchies). Third, participants must have learnt that taxonomic classification is favored among all possible classification rules and is the base of language (i.e., same names are used for same taxonomic category items and not other ways of classifying).

In summary, learning and using taxonomic concepts (e.g. for generalizing new words) requires a sufficiently developed and lexicalized conceptual building mechanisms. However, how do children acquire these sufficiently developed mechanisms for concept building? In the following section we shall review a conceptual development framework in which our empirical work can be read and discussed.

Conceptual development

Theories of conceptual development: finding the starting point of conceptual development

There is a controversy over the origins of conceptual development. Many theories argue for the existence of innate prior knowledge that guides early concept development. These views postulate that, from the first steps of development, top-down constraints are needed to guide relevant information selection. Such top-down processes are also necessary, in this view, for building concepts such as innate knowledge within “core” domains (Carey, 2009; Carey & Spelke, 1994; Spelke, 2000), skeletal principles (R. Gelman, 1990), ontological knowledge (Keil, 1992; Pinker, 1988), conceptual assumptions (Gelman, 1988; Markman, 1989), and word-learning biases (Golinkoff et al., 1994; E. M. Markman, 1989a).

Other theories consider that early in development, cognitive processes are grounded in powerful learning mechanisms, such as statistical and attentional learning (French et al., 2004; Mareschal et al., 2002; Sloutsky, 2003; Sloutsky & Fisher, 2004; Smith, 1989; Smith et al., 1996) rather than preexisting top-down mechanisms. In this theoretical framework, statistical and attentional learning mechanisms allow children to process perceptual input from the environment and build perceptual groupings, in a bottom-up manner. In this view, conceptual development then progresses from these early perceptual groupings, embryos of concepts, to highly abstract concepts.

These views’ standing points are opposite: conceptual development is constrained by top-down innate mechanisms, or by bottom-up, learning mechanisms. However, the main sticking point for all theories of conceptual development is the same: how are children’s cognitive processes constrained to build complex concepts. As mentioned above, the main challenge in concept building is selecting the relevant information to build the concept. In conceptual development, the aim is to understand how children’s cognitive processes develop to be able to select this relevant information.

Theories for prior knowledge, argue that the environment’s input is too impoverished (Poverty of the Stimuli argument, Chomsky, 1980; Cook, 1991; Gelman, 1990) to constrain concept building. Insufficiently developed attention mechanisms in young children may also lead children to attend to too much or wrong information leading them to fail in generalizing correctly what they have learned. In this view, limited selective attention mechanisms and impoverished input can’t account for the development of abstract rule-based categories characteristic of the complex human concept system.

General introduction

While the poverty of the stimuli argument exists, it also has its counterpart. Where some consider that the environment's input is too impoverished for guiding conceptual development, others on the contrary argue that the environment stimuli is highly regular sufficiently so to guide conceptual development even without prior knowledge (Sloutsky, 2010). This view is supported by neural evidence that shows that the neural system is sensitive to similarity (Goldstone & Son, 2012) and would thus be sensitive to the similarities in the environment as cues to understanding the world. Thus, children may learn language from the environment's input without prior "wired in" language capacities. Several lines of research have also argued that basic-level categories and broader-level categories have multiple within-category commonalities (French et al., 2004; Jones & Smith, 2002; Samuelson & Smith, 1999) frequent enough to direct attentional mechanisms, and thus to attract attention and entail learning. It can be kept in mind, that individual's sensitivity to similarity between stimuli is a prerequisite to the theory of alignment and the impact of comparison for concept learning that will be discussed further on.

In this attentional mechanism view, input and even limited attentional resources can account for the building of perceptual groupings (Sloutsky, 2010) or spatial image-schemas (Mandler, 2010). These first embryos of concepts likely group the environment's most salient features (i.e., by their perceptual salience, or their frequency in the environment, spatial information), which are the most likely to constrain attention through a bottom-up process from an early age. Traces of percepts - embryos of concepts - can be found in infants for example who recognize their mother's voice (Mills & Melhuish, 1974) and therefore have an embryo of concept built from this uni-modal stimuli.

Selective attention is therefore directed by the environments input's features. Considering the complexity and degree of abstractness of the human conceptual system, it is reasonable to question whether input from the environment can constrain attentional mechanisms strongly enough to build all of human concepts, and thus look for other ways attentional mechanisms may be constrained. However, we consider that one should not then turn to innate knowledge as an answer, but to constraints that are a product of prior acquired conceptual knowledge itself (Sloutsky, 2010). One can then consider that conceptual development is driven by the interplay between multiple elements: general cognitive processes (i.e., perception, memory systems, language, executive functions), more specific cognitive processes such as category learning mechanisms (Ashby et al., 1998), the nature of the environment's input (i.e., dense and sparse categories (Sloutsky, 2010), and learned conceptual knowledge (Murphy & Medin, 1985). As cognitive processes mature and learning processes come online, they can constrain selective attention better and allow more complex conceptual knowledge to develop. Reversely, as

conceptual knowledge develops, it also constrains selective attention. For example, category features learnt to be more important than others can be more heavily weighted than other less useful features (Hammer & Diesendruck, 2005; Sloutsky & Spino, 2004). Throughout the process of conceptual development, concept building can be considered a maturing process that combines an increasing number of sources of information. This interplay between newly discovered information and prior conceptual knowledge can be at play from very early stages of life. Indeed, as soon as even the simplest of uni-modal concepts are learned they can weigh features in newly encountered items and influence the building of new concepts. In other words, common features between the prior existing concepts and the newly encountered items will be weighted more heavily than others and thus more readily noticed and abstracted. A first step to constraining attention towards some features rather than others in a top-down way is at play.

From concept embryos to cross-modal concepts

As the development of attention processes (Lane & Pearson, 1982) and executive functions (Davidson et al., 2006; Diamond, 2013) enable children to consider multiple stimuli from the environment simultaneously, individuals become sensitive to regularities between items, which is a first necessary step towards classifying them together. Similarity in the environment will in turn help maintain attention to the environment's stimuli (Rothlein et al., 2018). It has been inferred, that at this stage in development, a category learning mechanism comes online that reduces and abstracts dense perceptual information from the environment to store a schematized representation in an implicit memory system (Ashby et al., 1998; Markman & Rein, 2013). Ashby et al., (1998) refers to this mechanisms as the compression-based category learning mechanism, Mandler (2010) as the Perceptual Meaning Analysis. Both the compression-based mechanism and implicit memory system are adapted for building grounded categories (Sloutsky, 2010). Thus, perceptually grounded cross-modal concepts, with multiple perceptual features compressed into a schematized representation and stored in implicit memory start to structure conceptual knowledge. From here, existing conceptual knowledge can no-doubtfully constrain future conceptual development.

At this stage, it is considered that concept building may encounter two limits. First, the attentional mechanisms described so far were attracted by the stimuli's salient properties in a bottom-up way. This is sufficient with perceptually dense categories in which category exemplars display numerous category defining features that stand out by their frequency. However, processes described so far may be insufficient for building perceptually sparse categories in which category exemplars display few

General introduction

category defining features (Kloos & Sloutsky, 2008). Second, the mechanisms we can so far account for, are sometimes considered to support perceptually grounded concept building and may not account for concepts including symbolic information such as linguistic content. For building sparser concepts and concepts with linguistic features, other learning mechanisms, explicit memory and a higher level of executive functioning must come online. This is also a precondition for concepts including linguistic information to emerge.

From noticing to selecting features

During the course of child development, as long as children's concepts are built from the features that stand out in the environment, many ways of categorizing the world's items may coexist. In other words, children may each follow different rules for categorizing the same items and these rules may change as a function of context. A child may sort her toys by color, or she may sort them as a function of their use (e.g., for bed, for playing outside, her's or her brother's etc). It is not unlikely that this way of sorting may also change as a function of context, for example the category "is mine" may greatly vary in size whether a child is asked to play with her toys or to tidy up her toys, most toys being hers in the first case and a lot less in the second...

There are so many ways of categorizing items because the environment is rich of information as to categorization rules and children progressively become aware of it. Items are repeatedly found in specific contexts that indicate sorting rules. Educational toys and books also stimulate children's item categorization along many rules: word learning books or sorting and stacking toys are good examples. Books display objects on pages by color, kind or theme. Sorting toys can stimulate color shape or size sorting. These multiple categorization cues are a fertile environment for bringing children to notice pertinent information and motivating them to differentiate between relevant and irrelevant information.

These categorization cues and stimulating activities from our social environment drive the next step of conceptual development: the coming on-line of top-down processes to constrain attention to relevant features. As children notice the categorization cues and complete categorization activities, they progressively develop top-down directed attention mechanisms that can be used later in conceptual development. The selection-based category learning mechanism (Ashby et al., 1998) that helps shift attention between irrelevant and relevant information is an example of such top-down mechanisms. Children's process at this stage are mature for learning specific categorization rules and building categories accordingly.

General introduction

From perceptual features to linguistic features

After the first year and a half or two years of child development, information processing mechanisms become sufficiently mature for children to notice and remember item-word co-occurrences. Children's "vocabulary spurt" during which their word learning and correct use of words thrive, is good evidence of their ability to remember item word associations (Gopnik & Meltzoff, 1987, 1993).

We can consider that children's word learning and conceptual development evolve in a tightly coupled manner (Borovsky & Elman, 2006; Gopnik & Meltzoff, 1987). Considering this, beginning of word learning may be considered as a turning point in their conceptual development. An interpretation could be that, as word-item co-occurrences are remembered the word must be stored in memory, with the item's representation, in other words as part of the item's concept. Integrating a name, that is of symbolic nature, as part of a concept is a first step towards abstract lexicalized knowledge. As these first pieces of lexical information integrate existing, perceptually grounded, concepts a turn in development is being operated between solely perceptual concepts and more abstract complex ones.

The integration of words in concepts has two other important implications. First, the use by others of words for the same conceptually related items indicates that there is a sorting rule to be followed to name items and communicate. Therefore, in children's environment, where sorting was very variable and context-dependent, language promotes taxonomic categorization above other categorization rules. This may also be referred to as the taxonomic assumption (Markman, 1990). Second, the existence of given conceptual categories, noticed by children when others use words, can help drive their search to understand what features support conceptual categories. From this point in conceptual development, children are on a path that will carry them to find nonperpetual, category-defining features and integrate these lexicalized features in their originally solely perceptual categories.

The language development course and conceptual development course continue. Sentences are acquired and can cue concept learning as well as integrate concepts as lexicalized pieces of knowledge. The main components of building concepts are put into place and the system can evolve towards more complex, abstract, and numerous concepts without suffering fundamental structural changes.

Learning new words for objects appears as a tipping point in conceptual development between perceptually grounded concepts randomly categorized and concepts with more abstract lexicalized features that follow set categorization rules. It is still hard to be sure of the exact nature of this step and

of its full importance, but it is most likely that it is one of the major steps in conceptual development and one that opens the way to complex and abstract human conceptual knowledge.

Given the importance of word learning in children's development, how children learn and generalize new words is a compelling question in cognitive science. Despite the considerable amount of work that has already addressed this question we set the task of investigating *how* children come to a conceptual generalization decision. Our aim was to investigate more thoroughly than it may have been done before, which strategies children use to generalize new words on a conceptual basis, what are these strategies' temporal dynamics and what processes may underlie them.

Investigating concept learning and generalization: novel noun generalization tasks.

For multiple reasons, investigating conceptual development and, more precisely, conceptual development during word learning is very challenging. Concepts develop over time, in interplay with the environment and ones' existing knowledge. This makes each individual's developmental course unique. For these reasons, among others, capturing conceptual development steps is difficult. When one focuses on the relations between word learning and conceptual development another difficulty is that the process develops at a young age (i.e. 18 to 36 months) when experimental designs must be carefully adapted to children's abilities. A widely used task for investigating concept learning and use in the case of word generalization is a novel noun generalization task, in 4- to 5-year-old children. This task, as we will discuss below, has the benefit of replicating concept construction and use. It also is with children that are old enough to complete such a task in various settings. However, this benefit is also a draw back. Indeed, 4- to 5-year-old children are much older than children that are really learning their first words and modeling their concepts accordingly. To avoid biases due to known and unknown items being processed in the same task items used to build the task are known by children, who are most often asked to learn and generalize pseudo-words.

When children learn new words for known items in their environment they must notice the word-item cooccurrence, map the item to its existing representation and abstract the word to integrate it to the existing concept. This process can be considered a re-representation process (Gentner & Colhoun, 2010a). This process is replicated in a novel word generalization task. Participants must understand the targeted concept from the task's items and must use the extracted concept to make an inference: the word generalization. The word generalization is an inference, participants infer which of the available

General introduction

choice options is the item that belongs to the same category as the learning items before generalizing the name. The generalization result - or chosen item - is the action that allows experimenters to understand what inference was made to make the generalization possible.

The task parameters can constrain the generalization performance. Many recent studies have shown the effect of comparison settings at learning that increases the proportion of taxonomic generalizations (Namy & Gentner, 2002; Price & Sandhofer, 2021). Comparison and word generalization tasks will be reviewed in Section 1.

In Section 2 and Section 3 we present empirical studies. A general line of reasoning that links the work is investigating how children come to a generalization decision, how they learn and then use a concept in this case to generalize a new word. We have approached this question in two different ways.

In Section 2, we investigate children's strategies in two word generalization tasks through the analysis of eye movements. Our aim is to describe children's generalization strategies during the novel noun generalization tasks. We focus on children's strategy profiles, their temporal dynamics, and their predictors of generalization success. We also follow up these studies by investigating which cognitive mechanisms may underlie the word generalization process.

In Section 3, we will change paradigm and study word generalization in a free choice design. In this way we will be able to study not only children's preferred generalization choice but the full scope of what they consider a generalization solution. These tasks give more detailed information about what children consider a generalization solution and information about how they come to a generalization decision through results in two different free choice settings and signal detection theory analyses.

References

- Anderson, J. R. (1990). *The Adaptive Character of Thought*. Psychology Press.
- APA Dictionary of Psychology*. (s. d.). Consulté 28 mars 2023, à l'adresse <https://dictionary.apa.org/>
- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442-481. <https://doi.org/10.1037/0033-295X.105.3.442>

- Ashby, F. G., & Maddox, W. T. (1993). Relations between Prototype, Exemplar, and Decision Bound Models of Categorization. *Journal of Mathematical Psychology*, 37(3), 372-400. <https://doi.org/10.1006/jmps.1993.1023>
- Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
- Borovsky, A., & Elman, J. (2006). Language input and semantic categories : A relation between cognition and early word learning. *Journal of Child Language*, 33(4), 759.
- Brooks, L. R. (1987). Decentralized control of categorization : The role of prior processing episodes. In *Concepts and conceptual development : Ecological and intellectual factors in categorization* (p. 141-174). Cambridge University Press.
- Carey, S. (2009). Where Our Number Concepts Come From. *The journal of philosophy*, 106, 220-254. <https://doi.org/10.5840/jphil2009106418>
- Carey, S., & Spelke, E. (1994). Domain-specific knowledge and conceptual change. In L. A. Hirschfeld & S. A. Gelman (Éds.), *Mapping the Mind : Domain Specificity in Cognition and Culture* (p. 169-200). Cambridge University Press. <https://doi.org/10.1017/CBO9780511752902.008>
- Chomsky, N. (1980). Rules and representations. *Behavioral and brain sciences*, 3(1), 1-15.
- Cook, V. J. (1991). The poverty-of-the-stimulus argument and multicompetence. *Interlanguage studies bulletin (Utrecht)*, 7(2), 103-117. <https://doi.org/10.1177/026765839100700203>
- Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years : Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, 44(11), 2037-2078. <https://doi.org/10.1016/j.neuropsychologia.2006.02.006>
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135-168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- French, R., Mareschal, D., Mermillod, M., & Quinn, P. (2004). The Role of Bottom-Up Processing in Perceptual Categorization by 3- to 4-Month-Old Infants : Simulations and Data. *Journal of experimental psychology. General*, 133, 382-397. <https://doi.org/10.1037/0096-3445.133.3.382>
- Gelman, R. (1990). First Principles Organize Attention to and Learning About Relevant Data : Number and the Animate-Inanimate Distinction as Examples. *Cognitive Science*, 14(1), 79-106. https://doi.org/10.1207/s15516709cog1401_5
- Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology*, 20(1), 65-95. [https://doi.org/10.1016/0010-0285\(88\)90025-4](https://doi.org/10.1016/0010-0285(88)90025-4)

- Gentner, D., & Colhoun, J. (2010). Analogical Processes in Human Thinking and Learning. In B. Glatzeder, V. Goel, & A. Müller (Éds.), *Towards a Theory of Thinking* (p. 35-48). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03129-8_3
- Gentner, D., & Namy, L. L. (2004). The Role of Comparison in Children's Early Word Learning. In *Weaving a lexicon* (p. 533-568). MIT Press.
- Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Éds.), *Perspectives on Language and Thought* (1^{re} éd., p. 225-277). Cambridge University Press. <https://doi.org/10.1017/CBO9780511983689.008>
- Goldstone, R. L., & Son, J. Y. (2012). *Similarity*. Oxford University Press.
- Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels : The case for a developmental lexical principles framework[*]. *Journal of Child Language*, 21(1), 125-155. <https://doi.org/10.1017/S0305000900008692>
- Gopnik, A., & Meltzoff, A. (1987). The development of categorization in the second year and its relation to other cognitive and linguistic developments. *Child Development*, 58(6), 1523-1531. <https://doi.org/10.2307/1130692>
- Hammer, R., & Diesendruck, G. (2005). The role of dimensional distinctiveness in children's and adults' artifact categorization. *Psychological Science*, 16(2), 137-144.
- Hull, C. L. (1920). Quantitative aspects of evolution of concepts : An experimental study. *Psychological Monographs*, 28, i-86. <https://doi.org/10.1037/h0093130>
- Jones, S. S., & Smith, L. B. (2002). *How Relevant properties children know the relevant properties for generalizing object names*.
- Keil, F. C. (1992). *Concepts, Kinds, and Cognitive Development*. MIT Press.
- Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds : Effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General*, 137(1), 52-72. <https://doi.org/10.1037/0096-3445.137.1.52>
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development*, 3(3), 299-321.
- Lane, D. M., & Pearson, D. A. (1982). The Development of Selective Attention. *Merrill-Palmer Quarterly*, 28(3), 317-337.
- Mandler, J. M. (2010). The spatial foundations of the conceptual system. *Language and Cognition*, 2(1), 21-44. <https://doi.org/10.1515/langcog.2010.002>

General introduction

- Mareschal, D., Quinn, P. C., & French, R. M. (2002). Asymmetric interference in 3- to 4-month-olds' sequential category learning. *Cognitive Science*, 26(3), 377-389.
https://doi.org/10.1207/s15516709cog2603_8
- Markman, A. B., & Rein, J. R. (2013). *The Nature of Mental Concepts*. Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780195376746.013.0021>
- Markman, E. M. (1989). *Categorization and naming in children : Problems of induction*. MIT Press.
- Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive science*, 14(1), 57-77.
- Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*. <https://doi.org/10.1037/0033-295x.85.3.207>
- Mervis, C. B. (1980). Category Structure and the Development of Categorization. In *Theoretical Issues in Reading Comprehension*. Routledge.
- Mervis, C. B., & Rosch, E. (1981). Categorization of Natural Objects. *Annual Review of Psychology*, 32(1), 89-115. <https://doi.org/10.1146/annurev.ps.32.020181.000513>
- Mills, M., & Melhuish, E. (1974). Recognition of mother's voice in early infancy. *Nature*, 252(5479), Article 5479. <https://doi.org/10.1038/252123a0>
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Murphy, G. L. (2016). Explaining the Basic-Level Concept Advantage in Infants... or Is It the Superordinate-Level Advantage? In *Psychology of Learning and Motivation* (Vol. 64, p. 57-92). Elsevier.
- Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition : Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(1), 70.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316. <https://doi.org/10.1037/0033-295X.92.3.289>
- Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears : Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, 131(1), 5.
- Pinker, M. D. S. (1988). Language learnability and language development. Cambridge MA: Harvard University Press, 1984. Pp. Xi + 435. *Journal of Child Language*, 15(1), 189-199.
<https://doi.org/10.1017/S0305000900012137>

General introduction

- Price, G. F., & Sandhofer, C. M. (2021). One versus many : Multiple examples in word learning. *Journal of Experimental Child Psychology*, 209, 105173. <https://doi.org/10.1016/j.jecp.2021.105173>
- Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3(3), 382-407. [https://doi.org/10.1016/0010-0285\(72\)90014-X](https://doi.org/10.1016/0010-0285(72)90014-X)
- Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 665-681. [https://doi.org/10.1016/S0022-5371\(75\)80055-7](https://doi.org/10.1016/S0022-5371(75)80055-7)
- Rips, L. J., Smith, E. E., & Medin, D. L. (2012). Concepts and categories : Memory, meaning, and metaphysics. In *The Oxford handbook of thinking and reasoning* (p. 177-209). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199734689.013.0011>
- Rosch, E. H. (1973). Natural categories. *Cognitive Psychology*, 4(3), 328-350. [https://doi.org/10.1016/0010-0285\(73\)90017-0](https://doi.org/10.1016/0010-0285(73)90017-0)
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382-439. [https://doi.org/10.1016/0010-0285\(76\)90013-X](https://doi.org/10.1016/0010-0285(76)90013-X)
- Rothlein, D., DeGutis, J., & Esterman, M. (2018). Attentional Fluctuations Influence the Neural Fidelity and Connectivity of Stimulus Representations. *Journal of Cognitive Neuroscience*, 30(9), 1209-1228. https://doi.org/10.1162/jocn_a_01306
- Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies : Do ontology, category structure and syntax correspond? *Cognition*, 73(1), 1-33. [https://doi.org/10.1016/S0010-0277\(99\)00034-7](https://doi.org/10.1016/S0010-0277(99)00034-7)
- Sloutsky, V. M. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences*, 7(6), 246-251. [https://doi.org/10.1016/S1364-6613\(03\)00109-8](https://doi.org/10.1016/S1364-6613(03)00109-8)
- Sloutsky, V. M. (2010). From perceptual categories to concepts : What develops? *Cognitive science*, 34(7), 1244-1286.
- Sloutsky, V. M., & Fisher, A. V. (2004). Induction and Categorization in Young Children : A Similarity-Based Model. *Journal of Experimental Psychology: General*, 133(2), 166-188. <https://doi.org/10.1037/0096-3445.133.2.166>
- Sloutsky, V. M., & Spino, M. A. (2004). Naive theory and transfer of learning : When less is more and more is less. *Psychonomic Bulletin & Review*, 11, 528-535. <https://doi.org/10.3758/BF03196606>
- Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological review*, 96(1), 125.

General introduction

- Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children : A dumb attentional mechanism? *Cognition*, *60*(2), 143-171.
- Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology*, *10*(4), 502-532.
- Spelke, E. S. (2000). Core knowledge. *American Psychologist*, *55*, 1233-1243.
<https://doi.org/10.1037/0003-066X.55.11.1233>
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*(4), 327-352.
<https://doi.org/10.1037/0033-295X.84.4.327>
- Waxman, S. R., & Hatch, T. (1992). Beyond the basics : Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language*, *19*(1), 153-166.
- Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations : Toddlers' and preschoolers' word-learning biases. *Child development*, *61*(5), 1461-1473.

SECTION 1: A THEORETICAL REVIEW

The following section contains the article: Lagarrigue Y., Stansbury, E., Thibaut, J-P, (submitted), Apprentissage et généralisation du lexique par comparaison : les mécanismes de l'expansion conceptuelle. *Année Psychologique*

Section 1: A theoretical review

Apprentissage et généralisation du lexique par comparaison : les mécanismes de l'expansion conceptuelle.

Résumé

De nombreuses études ont montré que l'apprentissage lexical bénéficie des situations de comparaison de stimuli par rapport à celles sans comparaison. Les situations de comparaison impliquent soit deux items associés au même mot et porteurs de la même dimension pertinente (comparaison intra) soit deux items différents sur la/les dimension(s) pertinente(s) (comparaison inter). Dans cette revue de questions, nous analysons le rôle de la comparaison pour trois types de mot : les noms, les verbes et les adjectifs. Ensuite, nous analysons les facteurs tels que la familiarité des stimuli ou la similarité entre les stimuli, sur les performances de généralisation. Enfin, nous montrons comment ces facteurs et les deux types de comparaison interagissent et influencent l'apprentissage des trois types de mots. Nous présentons également les travaux sur les stratégies de comparaison utilisées et sur la façon dont elles sont influencées par les paramètres de la tâche. Nous terminons par les limites des études disponibles et des ouvertures.

Mots-clés : Apprentissage, Généralisation, Comparaison, Lexique.

Abstract

Learning and generalization of the lexicon of adjectives, nouns and verbs by comparison: the mechanisms of conceptual expansion.

Numerous studies have shown the benefits of comparison situations, compared to non-comparison situations, in lexical learning tasks. In these studies, comparison can be distinguished into two categories: on the one hand the presentation of two items similar on the relevant dimension(s) (intra comparison) and on the other hand the presentation of two items different on the relevant dimension(s) (inter comparison). In this review, we'll give an overview of how comparison based learning can help generalizing three types of words : nouns, verbs and adjectives. Studies have also shown that several factors, such as familiarity with the stimuli or similarity between stimuli, influence the performance of generalizing a new word. In this theoretical review, we show how these factors and the two types of comparisons interact and influence the learning of three sort of words: nouns, adjectives and verbs. We also present a synthesis of the work on the comparison strategies and on how they are influenced by the task characteristics and conclude by the limits and openings given by the available studies today.

Key words : Learning, Generalization, Comparison, Lexica

Introduction

L'apprentissage d'un lexique et son utilisation adéquate est assurément une question ancienne et centrale de la psychologie cognitive du développement. Cette question repose sur deux piliers empirico/théoriques séparés, l'un phonologique l'autre conceptuel, ce dernier, qui est notre cible, interrogeant la base conceptuelle de l'apprentissage des nouveaux mots et leur généralisation à de nouvelles entités. L'architecture lexico-conceptuelle se construit progressivement au cours du développement et reflète la complexité des domaines conceptuels acquis. Elle a été et est l'objet de nombreuses études qui dépassent le cadre du présent article (voir Gelman, 2013; Mareschal, et al., 2010; Sloutsky, 2015 pour des synthèses). On sait que les premiers mots de l'enfant réfèrent à des entités de son environnement immédiat, à des actions quotidiennes (Nelson & Gruendel, 1981; Bassano et al., 2005), à des membres typiques plutôt que non typiques d'une catégorie, à des catégories de niveau de base (e.g. chien) plutôt que des catégories des niveaux subordonné (e.g., mammifère) ou superordonnant (e.g., labrador) (Murphy, 2002; Rosch, 1978), des objets plutôt que des relations (Gentner, et al., 2011) ou des verbes dont on connaît mieux aujourd'hui la complexité cognitive (Childers, 2016; Hirsh-Pasek & Golinkoff, 2000, pour une synthèse).

Notre revue théorique porte sur le rôle de la comparaison des stimuli durant l'apprentissage lexical et la structuration conceptuelle sous-jacente et, plus particulièrement, la comparaison de plusieurs stimuli présentés simultanément avec un même adjectif, un même nom, ou un même verbe. De nombreuses expériences montrent que la comparaison conduit à une généralisation lexicale plus efficace (donc proche de l'utilisation du mot correspondant par l'adulte) que la situation classique d'apprentissage souvent fondée sur un seul exemplaire d'apprentissage à la fois (Gentner & Namy, 2006). L'enjeu central d'un apprentissage lexical et conceptuel, dans les faits, réalisé avec un nombre limité de stimuli d'apprentissage est l'extension du mot cible à un nombre illimité de référents potentiels. Cet apprentissage lexical a un aspect quantitatif (apprendre plus de mots) mais aussi un aspect productif (généraliser correctement les mots appris) (Bassano, 2000; Clark, 1995, 2009; Hoff, 2013).

Les tâches d'apprentissage lexical sont souvent des tâches de généralisation qui veulent cerner les dimensions conceptuellement pertinentes extraites par l'enfant. On présente un objet, une scène, une interaction ou une action conjointement à la production d'un mot. Ensuite, on introduit plusieurs stimuli parmi lesquels l'enfant choisit celui ou ceux auxquels généraliser le mot. L'imagier pour enfant en est une illustration typique, qui combine l'image

Section 1: A theoretical review

d'un objet et la présentation de son nom, situation que l'on postule suffisante pour permettre une utilisation ultérieure adéquate de ce mot. Toutefois, généraliser peut s'avérer hasardeux car les stimuli catégorisés et dénommés sont multidimensionnels et chaque dimension ou combinaison de dimensions pourrait être a priori le support de la généralisation (voir le célèbre exemple de « gavagai » de Quine, 1960). Puisqu'on ne peut pas imaginer que le jeune enfant teste systématiquement toutes les dimensions cible possibles, les chercheurs ont postulé et testé l'existence de contraintes sur l'apprentissage lexical qui permettraient de limiter le nombre d'hypothèses à tester en les orientant vers certaines dimensions (voir Markman, 1994, pour une synthèse). Par exemple, l'enfant privilégie des objets de même forme plutôt que de même texture ou de même couleur (biais de la forme, - shape bias -, Landau et al., 1988; Kucker, et al., 2019, pour une synthèse), ou des objets de la même catégorie taxonomique plutôt que des objets liés thématiquement (voir Markman 1989). Ces contraintes souvent décrites ne sont pas l'objet de notre article.

Notre objectif central est l'analyse des situations de comparaison de stimuli et la manière dont elles contribuent à contraindre la signification des mots, leur extension, lorsque les dimensions pertinentes, qui devraient être le support de la généralisation, ne sont pas saillantes. En effet, les travaux récents, au-delà de l'apprentissage lexical, confirment que ces comparaisons sont propices à une généralisation fondée sur des similitudes conceptuelles plutôt que sur des similitudes perceptives de surface notamment lorsque les dimensions conceptuellement pertinentes sont peu saillantes ou ne sont pas celles que les participants regardent spontanément (Gentner & Namy, 1999; Kaminski & Sloutsky 2013).

Nous décrivons d'abord la situation de comparaison dans le cas de l'apprentissage lexical. Ensuite, nous synthétisons les travaux utilisant cette méthode de présentation pour trois types de catégories lexicales, les noms, les verbes et les adjectifs, afin d'établir la généralité du phénomène ainsi que les conditions dans lesquelles cet avantage se manifeste selon les types de mots. Nous poursuivons par une analyse des informations utilisées ou utilisables dans les situations de comparaison, et terminons en évoquons les variables psychologiques susceptibles d'influencer les résultats des activités de comparaison et l'organisation de la recherche de l'information durant ces comparaisons.

Des différences entre ces catégories de mots peuvent être anticipées puisque leurs contraintes cognitives ne sont pas identiques. Les noms d'objets réfèrent à des entités statiques, spatio-temporellement stables, ce qui permet d'analyser leurs caractéristiques dans la durée; les noms de relations réfèrent à une caractéristique, souvent invisible, liant des

Section 1: A theoretical review

entités alors que les adjectifs, notamment ceux qui s'appliquent aux objets, réfèrent à des propriétés particulières de ceux-ci et demandent que l'on s'écarte des autres propriétés des objets pour appréhender leur référence, alors que les verbes désignent des relations prédicatives, souvent entre deux arguments ou sans qu'un argument ne soit mentionné. En première approximation, les résultats montrent que les verbes apparaissent, en moyenne, plus tardivement dans le lexique que les noms (Childers & Tomasello, 2006).

Le paradigme de comparaison

Généraliser correctement un mot à de nouvelles situations demande la compréhension des dimensions partagées par les objets qui portent le même nom (par exemple les chaises). Ces similarités peuvent être perceptivement saillantes, comme la forme ou la couleur, ou au contraire peu saillantes mais importantes pour l'utilisation du mot correspondant à ces stimuli. Nous décrivons la situation de comparaison et quelles en sont les caractéristiques générales comparées à une situation d'apprentissage-généralisation lexicale fondée sur un seul stimulus comme dans les imagiers (i.e. une image, un objet, une action pour un nouveau mot). Partons d'un exemple avec les noms, la catégorie lexicale la plus étudiée dans le cadre de ce paradigme, pour l'adapter ensuite aux deux autres catégories lexicales, les verbes et les adjectifs. Nous définissons donc catégorie (ou catégoriel) au sens large d'un ensemble d'entités qui peuvent être regroupées sous un même mot. Par exemple, pour les fruits on applique le nom fruit à la catégorie des entités fruits, pour l'action de courir, on applique le verbe courir à la catégorie d'actions courir, et pour la propriété jaune, on applique l'adjectif jaune à la catégorie des phénomènes qui correspondent à la valeur de dimension couleur jaune.

Classiquement, pour étudier la généralisation de nom, on utilise un paradigme dans lequel on présente un seul stimulus auquel on attribue un pseudo-nom (« Ceci est un zatu. Regarde le bien. » pour une pomme) (voir Figure 1). Ensuite, dans la phase test (ou transfert), on présente plusieurs nouveaux stimuli et on demande au participant de sélectionner celui auquel on peut attribuer le nom attribué aux exemplaires d'apprentissage (« Regarde bien ces objets. Lequel d'entre eux est aussi un zatu ? »). Généralement, les items de transferts contrastent au moins deux options de réponse : un objet qui présente de nombreuses similitudes perceptives avec le stimulus d'entraînement mais n'appartient pas à la même catégorie taxonomique (e.g., une bougie ronde) et un autre objet qui ne présente pas (ou très

Section 1: A theoretical review

peu) de similitudes perceptives avec le stimulus de la première phase mais appartient à la même catégorie taxonomique (e.g., une banane de la catégorie « fruit »). Parfois, un troisième choix appartenant à la même catégorie thématique est utilisé pour dissocier les similitudes contextuelles des relations taxonomiques (e.g., du jus de pomme), (Figure 1). Le choix de l'enfant indique s'il base l'extension du nouveau mot sur des traits perceptuels, un lien thématique ou sur des traits conceptuels.

La situation « avec comparaison » est identique en tout point à ceci près que deux stimuli d'apprentissage, au moins, sont proposés. Deux types de comparaison peuvent être distingués, les comparaisons intra-catégorielles (dans le sens ci-dessus, l'ensemble des objets, des actions, ou des propriétés qui peuvent être lexicalisés avec le même mot) et les comparaisons inter-catégorielles (dans le sens ci-dessus, des objets, des actions, ou des propriétés, qui ne peuvent pas être lexicalisées avec le même mot). Dans les comparaisons intra, les deux objets/événements d'apprentissage comportent la ou les dimensions cible(s) et sont désignés par le même mot, pour que la comparaison rende les dimensions pertinentes plus saillantes. Dans les comparaisons inter (ou contraste), on présente deux objets/événements qui diffèrent sur la ou les dimensions cible(s) mais sont similaires sur les autres dimensions. Ces deux entités ne sont alors pas désignées par le même mot. Ces deux types de comparaison ont été l'objet d'études que nous présenterons ultérieurement (section 5). Des situations analogues sont utilisées pour les verbes, avec deux illustrations d'une même action désignée par le même verbe ou deux situations avec des caractéristiques communes mais une action différente n'étant pas désignées par le même verbe. Pour les adjectifs, deux objets différents mais illustrant la propriété cible et désignés par le même adjectif (« Les deux sont blickish. ») ou deux objets identiques excepté sur la caractéristique désignée par l'adjectif cible.

Les situations de comparaison/non-comparaison interagissent avec d'autres facteurs comme la distance sémantique entre les items comparés (deux pommes, ou deux fruits par exemple) ou entre les domaines de généralisation (d'autres pommes, d'autres fruits, ou d'autres aliments, dans le cas des noms), la présence ou non d'un mot (nom, verbe ou adjectif) ou pseudo-mot pour désigner la/les propriété(s) cibles (« Regarde celui-ci, c'est un zatu. » ou « Regarde cette chose-là. »).

La comparaison et les différentes catégories de mots : noms, verbes et adjectifs

La comparaison et la généralisation de noms

La contribution de Namy et Clepper (2010) est une excellente illustration des facteurs manipulés dans les études sur la comparaison. Cette étude sur des enfants de quatre ans contraste une condition sans comparaison et une condition avec comparaison avec deux objets. L'objectif de l'apprentissage d'un nom est la compréhension par l'enfant des similitudes, saillantes ou non, qui sous-tendent les classifications taxonomiques, c'est-à-dire celles qui rassemblent des objets de même nature plutôt que des similitudes perceptives superficielles (e.g., la couleur, dans bien des cas, est non pertinente, une voiture rouge et une tasse rouge; voir Gelman, 2003). Dans la condition avec comparaison, les auteurs ont également manipulé la présence d'un stimulus de contraste (ou inter-catégoriel, présent ou absent). Dans la condition avec comparaison « intra-catégorielle », on disait à l'enfant « Regarde celui-ci, c'est un zatu. » en introduisant le premier stimulus (une batte de baseball) et « Regarde celui-ci, c'est aussi un zatu. » en introduisant le deuxième stimulus (un club de golf). Dans la situation de comparaison inter-catégorielle (c'est à dire avec contraste), les auteurs introduisaient un item perceptivement semblable aux deux items d'apprentissage mais non membre de la catégorie (par exemple, une cravate) comme un non-zatu (« Regarde celui-ci, ce n'est pas un zatu »). Dans la condition sans comparaison, un seul stimulus était présenté, par exemple la batte de baseball introduite comme un zatu. Dans la phase de transfert, elles ont demandé aux enfants lequel, entre une raquette de tennis (choix taxonomique), un stylo (choix perceptif) et des chaussures de sport (choix thématique), était aussi un zatu (voir Figure 1). L'objectif des auteurs était de confronter l'efficacité des comparaisons intra et inter. Les résultats montrent que les enfants choisissent beaucoup plus souvent le choix taxonomique (la raquette de tennis, taxonomiquement proche mais -jugé-perceptivement éloigné) dans la situation de comparaison intra que dans les situations sans comparaison ou en comparaison inter. Ce résultat, répliqué à de nombreuses reprises (voir Gentner & Namy, 2006), se manifeste par un plus grand nombre de généralisations taxonomiques (correctes) que de généralisations perceptives (incorrectes) (voir Figure 1). Dans d'autres situations, la comparaison promeut la découverte de dimensions peu saillantes au détriment des dimensions saillantes mais conceptuellement non pertinentes (Graham et al., 2010; Augier & Thibaut 2013). Les auteurs ont également montré que la présence d'un contraste avait un effet moindre sinon absent. Ces effets de la comparaison ont été répliqués

Section 1: A theoretical review

pour des objets connus (Gentner & Namy, 1999), des parties d'objet (Gentner et al., 2007), et des objets non familiers (e.g., Graham et al., 2010; Augier & Thibaut, 2013; voir Lagarrigue & Thibaut, sous presse, pour une présentation de l'apprentissage des noms).

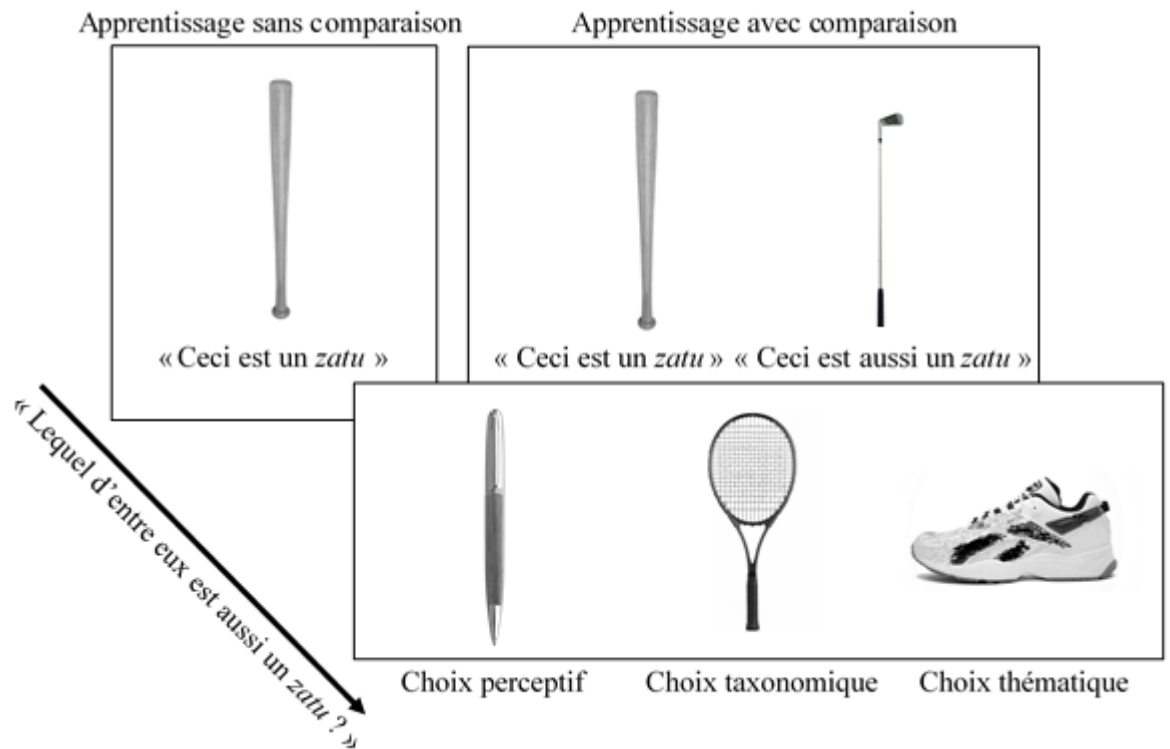


Figure 1 : Illustration d'une situation expérimentale de comparaison et de non-comparaison. // Illustration of an experimental situation of comparison and non-comparison.

Similarité entre les stimuli comparés

Dans les études sur la comparaison, les items d'apprentissage sont souvent proches perceptivement alors que les items test contrastent un item perceptivement proche mais non relié taxonomiquement et un item taxonomiquement relié mais différent perceptivement (voir Figure 1). Ce choix méthodologique s'explique par l'observation, très courante chez le jeune enfant, de l'influence des similitudes perceptives non pertinentes sur l'apprentissage et la généralisation, notamment la forme y compris lorsque ces similitudes perceptives sont non pertinentes (voir Laudau, 1994; Smith, 1989). Dans une représentation multidimensionnelle, les stimuli proches, sont aussi plus similaires que les éloignés (Nosofsky, 1986; Smith, 1989; Thibaut, 1997).

Section 1: A theoretical review

Objets familiers et objets non familiers

Dans la plupart des études, les auteurs utilisent du matériel connu et familier des enfants (un chapeau, un gâteau, un vélo ; voir Namy & Gentner, 2002). Les enfants peuvent donc s'appuyer sur des connaissances conceptuelles préalables pour réaliser la tâche. Présenter deux objets d'une même catégorie familière pourrait simplement aider à faire appel à des concepts déjà connus pour catégoriser les objets de transfert. A cette objection, les auteurs répondent que les items qu'ils demandent de comparer ne correspondent pas nécessairement à des catégories préexistantes dont le nom est connu par l'enfant. Par exemple, Gentner et Namy (1999) utilisent un vélo et un tricycle comme items d'apprentissage qui, en anglais, ne sont réunis dans aucune catégorie lexicale.

Les effets de la comparaison ont aussi été testés avec des stimuli non familiers, créés pour l'expérience, et dont les propriétés cibles conceptuellement pertinentes sont peu saillantes, comme c'est le cas chez Graham et al. (2010) et Augier et Thibaut (2013). Dans ces études, les stimuli sont inconnus des enfants. Ils sont construits autour de dimensions saillantes et conceptuellement non pertinentes telles que la forme, et de dimensions peu saillantes et conceptuellement pertinentes telles que la texture. Les résultats obtenus par Lagarrigue, Augier et Thibaut (en préparation) ont confirmé ceux obtenus dans les situations précédentes. Dans les situations sans comparaison, les généralisations étaient basées sur la dimension saillante mais non pertinente, la forme, alors que dans les conditions sans comparaison les généralisations étaient basées sur la dimension pertinente mais peu saillante, la texture. Une situation de comparaison permet, par alignement des stimuli, l'émergence de dimensions qui ne seraient sinon pas prises en compte sans ces comparaisons.

Les noms relationnels

Les noms que nous avons mentionné jusqu'à présent réfèrent à des catégories d'objets dont les propriétés sont visibles et spatio-temporellement stables. Les noms qui désignent des relations entre les objets plutôt que les objets eux-mêmes, comme « voisin » ou « addition » ou « ami », « ennemi » ou encore « combat » sont différents et plus difficiles à comprendre et à interpréter par les jeunes enfants car ils réfèrent à une relation entre des entités plutôt qu'aux entités elles-mêmes. L'addition désigne ainsi une relation entre des quantités, quelles que soient ces quantités. Des entités proches l'une de l'autre sont « voisines » quelle que soit leur nature. La comparaison peut se révéler prépondérante pour l'apprentissage lexical relationnel comme le montrent les études de Kotovsky et Gentner (1996), Gentner et al. (2011) ou de Thibaut et Witt (2015). Ces auteurs ont présenté deux paires d'objets familiers illustrant une

Section 1: A theoretical review

même relation à laquelle ils ont attribué un pseudo-mot. Par exemple, pour un couteau et une pastèque, ils indiquaient « Le couteau est le dax de la pastèque. », suivis par un autre couteau et une orange « Le couteau est aussi le dax de l'orange. ». Au test, ils présentaient un objet, par exemple une feuille de papier, et demandaient aux enfants parmi trois nouveaux objets, celui qui était le dax de de la feuille. Le choix relationnel illustre la même relation que les exemples (ici une paire de ciseaux), le choix thématique était relié thématiquement à la feuille de papier (un crayon), et le choix taxonomique appartenait à une même catégorie que la feuille de papier (un bloc note). Les situations avec comparaison et utilisation du langage donnent beaucoup plus de choix relationnels que des situations sans comparaison ou sans utilisation du langage. Ces résultats indiquent que la comparaison ne permet pas seulement la sélection d'objets sur une base perceptive, mais aussi des propriétés du réel moins immédiates, comme les relations entre les objets.

La comparaison et la généralisation de verbes

La classe lexicale des verbes réfère à des actions. D'abord, le verbe ne s'utilise pas en dehors d'une structure prédicative à un argument au moins comme « Le bébé mange ; la boîte tombe » et est « prédié » au sujet d'autre chose, par exemple, montrer un oiseau à l'enfant en disant « L'oiseau vole ». Pour comprendre la référence de vole, l'enfant doit comprendre que ce mot fait référence à une propriété (l'action) attribuée à l'oiseau. Or, les propriétés qui sous-tendent la référence des verbes sont souvent instables et peu visibles (un oiseau ne vole pas en permanence). Les verbes partagent donc, en partie, les difficultés associées aux noms de relation (voir ci-dessus). Comme eux, le verbe qualifie une relation entre deux arguments (un acteur et un patient) et la « mise en relation » peut être effectuée par des acteurs très différents et dans des contextes différents. Par exemple pour le verbe ouvrir, partant de « Lucie ouvre la porte », on peut modifier l'acteur « Marius ouvre la porte », l'objet « Lucie ouvre la fenêtre » voire le type d'ouverture (reposant sur des mouvements différents) « Lucie ouvre un pot de confiture ». Les contextes de généralisation du verbe ouvrir peuvent être très différentes de la/des situation(s) d'apprentissage.

Les verbes sont souvent décrits comme plus difficiles à apprendre que les noms (Bornstein et al., 2004; Gentner, 1982; Tardif et al., 1997). Ils apparaissent plus tardivement que les noms dans le lexique des enfants, vers 20 mois. Waxman et al. (2009) ont montré que les enfants de 24 mois généralisent incorrectement les verbes mais savent qu'un verbe fait référence à un événement alors qu'un nom fait référence à un objet.

Section 1: A theoretical review

Pour confronter des conditions avec et sans comparaison, on utilise un paradigme semblable à celui utilisé pour l'étude des noms, avec une ou plusieurs scènes dynamiques dans lesquelles apparaît un événement associé à un nouveau verbe (souvent créé pour la tâche dit « pseudo-verbe »). On demande ensuite aux participants de généraliser le verbe à un événement qu'ils doivent choisir parmi plusieurs événements proposés. La similarité entre les événements présentés au test et les événements d'apprentissage varie sur une ou plusieurs des dimensions de l'événement (les acteurs impliqués, l'objet utilisé pour l'action, l'action elle-même, le résultat de l'action, etc.). Le choix de l'enfant indique s'il généralise le verbe à une situation identique ou une situation différente et les dimensions utilisées pour généraliser (e.g., Childers, 2008).

Avant d'envisager les différentes dimensions qui influencent l'apprentissage de la référence des verbes, il faut noter que l'introduction du verbe (e.g., regarde le chien daxe" durant la présentation de l'action donne de meilleurs résultats lors de la généralisation qu'une présentation sans cette lexicalisation (e.g., regarde le chien, tu vois ce qu'il fait). Ce résultat analogue à celui décrit pour les noms (voir ci-dessus) a été reproduit pour les verbes par Childers (2011).

A l'instar des noms, les enfants doivent dépasser les similitudes perceptives non pertinentes pour généraliser les verbes et négliger les éléments saillants mais conceptuellement non pertinents au profit des propriétés pertinentes de l'action. Dans leur cas, on a évoqué un biais de conservation (Tomasello, 1992). Tomasello a montré qu'au début de l'acquisition du lexique, les jeunes enfants généralisent les verbes à des événements très similaires à l'événement d'entraînement et ont des difficultés lorsque les situations d'apprentissage et de test diffèrent sur plusieurs dimensions. L'apprentissage d'un verbe à partir d'une seule occurrence d'un événement ont souvent révélé des généralisations à des événements similaires perceptivement mais incorrects (Behrend, 1995; Gropen et al., 1991). Ces choix fondés perceptivement diminuent lorsque plusieurs exemplaires de la même action sont comparés pendant l'apprentissage du verbe (Childers 2008; Childers & Paik 2009; Forbes & Farrar, 1995; Scott & Fisher, 2009). Ces représentations multiples de « l'action » permettent aux enfants d'en élaborer une représentation permettant à la fois la mise en avant des propriétés conceptuelles pertinentes et l'intégration des variations possibles des propriétés contextuelles. D'une part, la proximité entre les situations d'apprentissage permet d'attirer l'attention des enfants sur l'événement lui-même et de repérer les aspects conceptuels identiques d'une situation à l'autre. D'autre part, la variabilité des événements permettrait

Section 1: A theoretical review

l'élaboration d'une représentation de la variation possible des dimensions non pertinentes et la généralisation du verbe à de nouvelles situations. Nous examinons maintenant comment les similarités entre les items d'apprentissage (la présence d'un même mot ou l'utilisation d'un même objet pour réaliser l'action) facilitent la comparaison et la généralisation des verbes et, ensuite, l'impact de la variabilité des situations.

La similarité entre les situations d'apprentissage comparées

La similarité des situations d'apprentissage des verbes est au centre de l'étude de Haryu et al. (2011). En manipulant cette similarité, les auteurs montrent que présenter deux objets similaires pour réaliser l'action entre les événements d'apprentissage et l'événement présenté au test facilite la généralisation du verbe. Un premier groupe, qui a obtenu les meilleures performances de généralisation, comparait deux occurrences d'une même action exécutées avec des objets de même taille et de même forme. Un second groupe d'enfants comparait deux exemples d'une même action où les objets différaient par la taille et la forme. Les actions sont complexes et multidimensionnelles (un acteur, l'action, le patient, l'objet, etc.). Les similitudes attirent l'attention sur des points communs, comme c'est le cas pour les noms, et permettent d'initier le processus de comparaison qui, sans cela, ne pourrait débiter car les enfants ne comprendraient pas par quelle dimension commencer. Toutefois, une trop grande similarité entre les événements à comparer peut « annuler » l'effet positif de la comparaison de plusieurs événements. Ainsi, dans l'étude de Namy et al. (2007), les enfants ayant vu des événements réalisés avec objets identiques ont de moins bonnes performances de généralisation que des enfants ayant vu des événements réalisés avec objets similaires.

La variabilité des événements à comparer

Nous avons ci-dessus que, dans le cas des noms, leur généralisation à des objets conceptuellement distants est favorisée par l'augmentation de la distance conceptuelle entre les objets utilisée à l'apprentissage (Gentner et al. 2011; Thibaut & Witt 2015, 2017). Pour les verbes, augmenter la différence entre les événements d'apprentissage favorise également la généralisation à des situations plus lointaines, à condition que la difficulté générée par les variations des paramètres de la tâche reste dans les limites des capacités de contrôle des sujets (e.g., Augier & Thibaut, 2013). Pour manipuler la différence entre les événements à comparer dans une tâche de comparaison-généralisation, on modifie une dimension de l'événement, l'acteur, le patient ou l'action, d'un exemple à l'autre. Par exemple, le verbe soulever désigne la même action d'un exemple à l'autre alors que les acteurs et patients diffèrent. Pour un verbe intransitif, comme rouler, le mouvement pourra se faire sur un plan plat ou incliné

Section 1: A theoretical review

(dimension support) ou résulter en une chute ou un passage derrière une paroi (dimension résultat).

Si la diversité des événements influence la généralisation des verbes, elle permet aussi l'élaboration des représentations détachées du contexte d'apprentissage. Childers (2011) le montre avec une tâche de comparaison de deux événements dans laquelle sont manipulés l'action ou le résultat. L'événement initial est une balle qui roule sur le couvercle incliné d'une boîte avant d'y tomber et y être cachée (action : rouler, résultat : être caché). Selon la condition expérimentale, l'événement utilisé pour la comparaison est soit identique à l'événement initial (condition contrôle) soit diffère sur l'action réalisée (la balle est posée dans la boîte, ou mise dans un sac, ou recouverte selon le cas, condition résultat), soit diffère sur le résultat (la balle roule mais reste visible à la fin de l'essai, condition action). Durant ces manipulations, les différentes occurrences d'un événement sont toujours désignées par le même verbe.

La tâche de généralisation est une tâche de reproduction de l'action. Les enfants testés, âgés de 30 mois, reproduisent mieux l'action cible lorsque les événements d'apprentissage comparés varient sur une dimension (le résultat ou l'action) que si les événements d'apprentissage sont identiques (condition contrôle). Dans ce dernier cas, les enfants reproduisent les comportements vus durant l'apprentissage de manière strictement identique et ne généralisent pas le verbe appris à des événements différents de ceux vus lors de la comparaison. Or, c'est la généralisation du verbe à d'autres événements désignés mais différents des événements ayant permis l'apprentissage du verbe qui est recherché. Quand une dimension des événements comparés (le résultat ou l'action) change d'un exemple à l'autre, les enfants généralisent à la fois le verbe à des actions identiques de ceux vus au moment de la comparaison et à des événements différents. C'est donc la variation sur une dimension des événements comparés qui permet de généraliser le verbe à des événements différents de ceux vus durant l'apprentissage de ce verbe.

D'autres études ont porté sur l'effet de la variabilité des événements d'apprentissage en faisant varier d'autres dimensions que celles manipulées par Childers (2011), telles que l'acteur de l'action (Childers, 2017; Scott & Fisher, 2012), l'action et l'objet de l'action (Snape & Krott, 2018), l'outil utilisé pour exercer l'action (Childers, 2019) et l'objet sur lequel s'exerce l'action (Childers, 2020). Cependant, certaines études montrent que ces variations font parfois baisser les performances des enfants. Par exemple, Maguire et al. (2008), ont montré que les enfants de deux ans et demi ont plus de difficultés à apprendre un

Section 1: A theoretical review

nouveau verbe lorsqu'on leur montre deux événements d'apprentissage dans lesquelles les acteurs varient que lorsqu'on leur montre deux événements avec le même acteur. Dans cette étude, la moitié des enfants voyaient quatre acteurs différents réaliser une action pendant la phase d'apprentissage et l'autre moitié voyaient le même acteur réaliser quatre fois la même action. Au test, les enfants voyaient deux nouvelles actrices réaliser soit l'action vue pendant l'apprentissage, soit réaliser une autre action. Les résultats montrent que les réponses des enfants étaient au-dessus du hasard dans les deux conditions mais significativement meilleures dans la condition avec une seule actrice pendant l'apprentissage.

En conclusion, les études sur l'apprentissage des verbes bénéficient de la comparaison. A l'intérieur de ce paradigme, la distance sémantique entre les événements et la variabilité de la mise en scène sont des facteurs qui vont moduler les effets de la comparaison. En somme, des effets analogues à ceux trouvés pour les noms.

La comparaison et la généralisation d'adjectifs

Les adjectifs réfèrent à des propriétés des objets qui n'existent pas indépendamment des objets qui les portent (e.g., jaune, sale). Ils renvoient à des propriétés observables (noir, cassé), transitionnelles liées aux émotions et à des états physiologiques (content, fatigué), ou à des traits stables (gentil, sage) (Graham et al., 2005). De manière analogue aux noms et aux verbes, les situations de comparaison permettent une meilleure généralisation des adjectifs que les situations de non-comparaison (Hall et al., 1993; Mintz & Gleitman, 2002).

En situation de comparaison, Waxman et Klibanoff (2000) ont montré que présenter plusieurs objets similaires en tous points excepté sur la propriété à généraliser (par exemple une chaise verte et la même chaise mais bleue) ou deux objets différents en tous points excepté sur la propriété à généraliser (un verre transparent et une assiette transparente) facilite la détection de la propriété cible. Dans une première expérience, un des objets à comparer portait une propriété (par exemple transparent) et l'autre objet d'apprentissage était identique sauf qu'il avait la valeur « inverse » (opaque) de cette même propriété. Dans une des deux conditions comparées, les deux items d'apprentissage étaient membres de la même catégorie de base (une assiette transparente et une autre assiette opaque) et dans l'autre condition, ils étaient membres de catégories de base différentes (une assiette transparente et une brosse à dent opaque). Les objets étaient introduits avec un pseudo-adjectif (« Regarde celui-ci, il est zatou. », « Regarde, celui-là aussi il est zatou »). Au test, les enfants devaient choisir entre deux objets d'une autre catégorie (une bouteille transparente et une bouteille opaque) celui

Section 1: A theoretical review

auquel on pouvait attribuer l'adjectif cible (« C'est lequel qui est aussi zatou ? ») présentant la propriété cible. Les résultats montrent que présenter deux exemples de la même catégorie de base aide les enfants de trois ans à généraliser le nouvel adjectif sur la base de la propriété cible, y compris lorsque les objets tests appartiennent à des catégories de base différentes des objets d'apprentissage. En revanche, conformément aux hypothèses des auteurs, les conditions « deux exemples de catégories de base différentes » et « sans comparaison » ne facilitent pas la généralisation.

Dans la seconde expérience, Waxman et Klibanoff (2000) présentaient deux items différents de la même catégorie (par exemple un insecte rouge brillant et un insecte bleu brillant) mais identiques sur la propriété cible (brillant). Les enfants devaient ensuite choisir entre deux objets (un hippopotame rouge brillant et un hippopotame bleu mat) celui qui présentait la propriété cible. Les auteurs ont également manipulé le niveau de catégorisation des items d'apprentissage, soit de la même catégorie de base, soit appartenant à deux catégories de base différentes (un insecte rouge brillant et un chien bleu brillant). Les résultats montrent de meilleurs résultats lorsque les deux exemples sont issus de la même catégorie de base que lorsqu'ils sont issus de deux catégories de base différentes. Cette étude a montré que les deux types de comparaison, intra et inter, sont efficaces pour l'apprentissage d'adjectifs. La seconde précise ce résultat en montrant que la variabilité ne doit pas être trop importante (ne doit pas porter sur de nombreuses dimensions) pour que enfants puissent saisir à quelle dimension réfère l'adjectif.

Dans l'étude précédente, les résultats montrent que la comparaison intra ou inter de deux objets qui n'appartiennent pas à la même catégorie de base n'aide pas les enfants à généraliser la propriété cible sur un nouvel objet. Cependant, Booth et Waxman (2003) ont montré que si les deux types de comparaison (intra et inter) sont présents conjointement durant l'apprentissage, les enfants généralisent le nouvel adjectif sur la propriété cible quel que soit le niveau de catégorisation qui unit les objets d'apprentissage. Pour cela, elles ont utilisé un protocole dans lequel elles montraient quatre exemples de la même catégorie présentant la dimension cible (par exemple ils étaient tous de la même couleur). Dans une des conditions, les quatre exemples faisaient partie de la même catégorie de base (e.g., quatre chevaux marrons), dans une seconde condition, ils faisaient partie de la même catégorie superordonnante (e.g., quatre animaux différents marrons). Ensuite un autre item d'apprentissage issu d'une autre catégorie superordonnante (e.g., une pomme) et présentant une propriété contrastive (e.g., rouge) était introduit. Tous les objets étaient présentés avec un

pseudo-mot. Au test, les enfants devaient choisir entre un objet de la même catégorie que les items d'apprentissage (un autre animal bleu) ou un objet présentant une propriété commune avec les items d'apprentissage (une chaise marron). Leurs résultats montrent que quel que soit le niveau de catégorisation des objets présentés à l'apprentissage, les enfants généralisent le nouvel adjectif sur l'objet qui porte la propriété cible.

Les différents types de comparaison

Dans les expériences que nous avons présentées pour les différents types de mots, la comparaison peut impliquer deux types de stimuli, d'une part, les items d'apprentissage introduits par un mot (nom, verbe, adjectif) désignant des entités, des phénomènes, d'une même catégorie d'entités (objets, relations, actions, caractéristiques d'un objet) que nous avons appelés *intra*, et d'autre part, les items d'apprentissage introduits comme des items d'une autre catégorie d'entités, *contrastive*, et que nous avons appelés *inter*. Cette section débute par une analyse systématique de la contribution de ces deux types de comparaisons durant l'apprentissage de nouveaux noms que nous étendons ensuite aux adjectifs et aux verbes.

Définition des comparaisons intra et inter

Dans les comparaisons *intra*, on présente deux objets/événements d'apprentissage qui comportent la ou les dimensions cible(s) pour la/les rendre plus saillante(s). La comparaison montre que les caractéristiques communes entre les stimuli peuvent être des critères diagnostiques potentiels. Dans les comparaisons *inter* (ou *contraste*), on présente deux objets qui diffèrent sur la ou les dimensions cible(s) mais sont similaires sur les autres dimensions en les classant dans des « catégories » d'entités différentes. Les caractéristiques communes aux stimuli contrastés sont donc non diagnostiques car elles ne permettent pas de distinguer les catégories. A l'inverse, leurs différences deviennent des critères diagnostiques potentiels. On peut introduire un contraste en présentant le deuxième objet comme un non-membre de la catégorie du premier (« Celui-ci n'est pas un zatu. », « celui-ci n'est pas blickish »), soit en présentant le deuxième objet comme membre d'une autre catégorie (« Celui-ci est un zatu et celui-là est un duma. », « celui-ci est blickish et celui-là est daxish ») (Childers et al., 2014; Hall, & Rhemtulla, 2014). Le tableau I présente les situations de non-comparaison et de comparaison, ces dernières dans les comparaisons *intra* et *inter*, pour les trois types de mots étudiés.

Section 1: A theoretical review

Tableau I : Illustrations des différentes conditions de comparaison pour les trois types de mots étudiés. // Illustrations of the different comparison conditions for the three types of word

Type de comparaison		Nom (un Zatu = « Pomme »)		Verbe (zatuser = « Lancer »)		Adjectif (C'est zatusé = Jaune »)	
		Ce que l'on dit	Ce que l'on montre	Ce que l'on dit	Ce que l'on montre	Ce que l'on dit	Ce que l'on montre
Aucune	Item 1	Ceci est un <i>zatu</i>	Une pomme Golden	Il zatuser	Paul qui lance un ballon	C'est <i>zatusé</i>	Une pomme Golden
Intra-catégorie	Item 1	C'est un <i>zatu</i>	Une pomme Golden	Il zatuser	Paul qui lance un ballon	C'est <i>zatusé</i>	Une pomme Golden
	Item 2	Lui aussi c'est un <i>zatu</i>	Une poire	Lui aussi il <i>zatuser</i>	Pierre qui lance un caillou	Celui-ci aussi est <i>zatusé</i>	Une poire
Inter-catégorie	Item 1	C'est un <i>zatu</i>	Une pomme Golden	Il zatuser	Paul qui lance un ballon	C'est <i>zatusé</i>	Une pomme Golden
	Item 2	Celui-ci ce n'est pas un <i>zatu</i> // Celui-ci est un <i>dax</i>	Une balle de tennis	Lui il ne <i>zatuser</i> pas. // Lui il <i>daxise</i>	Paul qui pose un ballon	Celui-ci n'est pas <i>zatusé</i> // Celui-ci est <i>daxé</i>	Une pomme Granny Smith

Hammer et collègues ont tenté de mesurer la quantité d'information apportée par chaque type de comparaison, intra et inter. Par exemple, Hammer et al. (2009) ont créé des stimuli (des créatures multidimensionnelles) regroupés en catégories sur la base de quatre caractéristiques binaires : la couleur des yeux (bleu ou jaune), la forme de la queue (droite ou tordue), la couleur de la fourrure (jaune ou verte) et la couleur de la peau (bleu ou orange).

Section 1: A theoretical review

Dans la phase d'entraînement, deux créatures étaient présentées simultanément. En fonction de la condition expérimentale on indiquait soit qu'elles appartenaient à la même catégorie (condition intra), soit qu'elles faisaient partie de deux catégories différentes (condition inter). Dans la phase test, les participants devaient décider si les deux créatures faisaient partie de la même catégorie ou non. Les auteurs ont contrôlé la quantité d'information délivrée par chacun de ces deux types de comparaison en répertoriant toutes les hypothèses possibles (seize) de catégorisation et en comptant le nombre d'hypothèses éliminées par chaque type de comparaison. Ils ont mis en place des situations de comparaison inter et intra qui permettaient d'éliminer le même nombre d'hypothèses et ont comparé leurs effets réels. Leurs résultats montrent que tous les groupes d'âge ont correctement catégorisé les créatures pendant la phase test lorsque l'apprentissage était réalisé en comparaison intra. Les adultes et les enfants de dix ans ont correctement classé les tests en condition inter alors que les enfants de six ans ont répondu au hasard. Puisque les deux types de comparaison étaient équivalentes, Hammer et al. (2009) en ont conclu que l'utilisation des dimensions de contraste comme non pertinentes n'est pas mature avant dix ans. Des études récentes ont répliqué ce résultat d'une influence faible des items de contraste (Ankowski et al., 2013; Augier & Thibaut, 2013). Selon Augier et Thibaut (2013) la difficulté des situations de contraste proviendrait de la nécessité de considérer positivement les dimensions communes intra tout en négligeant des similitudes inter, donc en les inhibant. En d'autres termes, certaines similitudes doivent être valorisées, d'autres négligées, simultanément.

Hammer et al. (2008) se sont intéressés à des situations plus écologiques pour comprendre comment les propriétés des objets présentés influencent les effets de la comparaison. Selon eux, en comparaison intra, plus les objets présentés sont différents, plus le processus de comparaison est informatif. Par exemple, si on veut faire apprendre le mot fruit à un enfant, lui montrer une banane et un ananas, deux fruits très différents, sera très informatif car seules les dimensions communes entre les deux (le fait qu'ils soient des végétaux, comestibles, sucrés) sont pertinentes pour catégoriser un nouvel élément dans cette catégorie. Complémentairement, les différences telles que la couleur, la taille ou la forme pourront être jugées non pertinentes pour l'appartenance. Autrement dit, en intra-catégoriel, plus le nombre de dimensions sur lesquelles diffèrent les objets présentés augmente, plus l'informativité de la comparaison augmente. C'est pour cette raison que présenter deux objets identiques en comparaison intra s'avère peu efficace (Namy et al., 2007).

Section 1: A theoretical review

A l'inverse, dans une situation de comparaison inter (ou contraste) plus les objets présentés et contrastés se ressemblent, plus la situation est informative. Ainsi, présenter un fruit et un non-fruit très ressemblant, comme une pomme rouge et un ballon rouge, est très informatif car cela permet d'exclure toutes les dimensions communes (la couleur, la forme, la taille) des dimensions diagnostiques. En effet, être rond, rouge et tenir dans la main ne permet pas de dire s'il s'agit d'un fruit et que ce sont d'autres caractéristiques de la pomme qui font d'elle un fruit. On notera que savoir que certaines dimensions ne sont pas diagnostiques ne nous dit pas lesquelles, parmi les dimensions restantes, le sont.

Selon Hammer et al. (2008), le processus de recherche de différences, c'est à dire les situations de comparaison inter, serait particulièrement efficace pour apprendre les limites entre les catégories et est donc plus utile pour l'apprentissage des catégories très ressemblantes, comme les catégories de niveau subordonné. Pour distinguer un labrador d'un berger allemand, les informations du niveau de base (i.e., chien, ils aboient, ils ont quatre pattes) ne sont pas informatives. Il faut se référer à des propriétés spécifiques telles que les labradors ont des oreilles pendantes, pas les bergers allemands pour distinguer les deux races l'une de l'autre. Ainsi, c'est la comparaison des deux races de chien qui permet de noter les différences entre les deux et donc de découvrir les dimensions qui les distinguent (voir Rosch, 1978). Cette logique a été utilisée dans des apprentissages professionnels, notamment en médecine radiologique. La différence entre un tissu sain et une tumeur apparaîtra plus clairement de la confrontation des deux (inter) que de la répétition de l'un ou de l'autre (intra) (Boucheix, et al., 2020 pour une présentation; Kok et al., 2013).

Les comparaisons intra et inter dans le cas des verbes et des adjectifs

L'analyse proposée par Hammer et ses collègues aux noms d'objets peut également s'appliquer aux adjectifs et aux verbes. Le nouveau mot ne fait alors plus référence à l'objet lui-même mais à une de ses propriétés pour les adjectifs et à une action pour les verbes. La catégorie d'appartenance des objets impliqués ne permet pas de distinguer les deux types de comparaison dans le cas des objets et des verbes. Dans ces derniers cas, on parle de comparaison intra lorsque les deux items d'apprentissage présentent la/les propriété(s) cible(s) et peuvent être désignés par le même mot comme, par exemple, une pomme jaune et une banane pour l'adjectif jaune, ou un lancer de balle et un lancer de javelot pour le verbe lancer. Comme pour les noms, l'enfant doit déduire que les propriétés qui diffèrent entre les items d'entraînement ne sont pas celles qui sont désignées par le nouvel adjectif mais que celui-ci réfère plutôt à leur(s) point(s) commun(s). On retrouve donc le principe de Hammer et al.

Section 1: A theoretical review

(2008) sur les comparaisons intra : plus les objets diffèrent sur un grand nombre de dimensions et plus la comparaison est informative. C'est la même chose pour les verbes, plus le nombre de dimensions sur lesquelles les événements d'apprentissage diffèrent est élevé, plus la comparaison est informative. Les études montrent que les situations de comparaison intra sont efficaces pour l'apprentissage d'adjectifs et de verbes (voir par exemple Waxman & Klibanoff, 2000; Childers, 2011) dans le cadre de certaines limites. Dans le cas des adjectifs, Waxman et Klibanoff (2000) ont montré que si les deux items d'entraînement ne font pas partie de la même catégorie de base, les enfants de trois ans généralisent incorrectement le nouveau mot.

Dans le cas des comparaisons inter, par exemple on peut faire apprendre l'adjectif jaune en présentant deux items qui diffèrent sur cette propriété, une pomme jaune et une pomme rouge ou le verbe « lancer » en montrant un personnage qui lance une balle et le même personnage qui pose la même balle. On constate également, comme l'ont suggéré par Hammer et al. (2008), que l'augmentation du nombre de dimensions sur lesquelles les objets diffèrent diminue l'informativité de la comparaison. A la différence des noms, les études révèlent une efficacité des comparaisons inter pour l'apprentissage des adjectifs et des verbes (voir Waxman & Klibanoff, 2000; Childers, 2011). A ce stade, on ne peut cependant pas conclure, comme pour les noms, à la supériorité d'un type de comparaison sur l'autre car ils n'ont pas été directement opposés.

Influence des paramètres des tâches de comparaison : mécanismes et stratégies en jeu

Nous considérons maintenant les paramètres pouvant influencer les comparaisons pour obtenir une meilleure compréhension des mécanismes cognitifs et des stratégies qui les soutiennent. Quand la manipulation d'un paramètre d'une tâche de comparaison-généralisation modifie les performances des sujets, l'interprétation des différences observées permet de faire des hypothèses sur les mécanismes cognitifs en jeu et/ou les stratégies de résolution utilisées.

Les mécanismes en jeu

L'effet du nombre d'items à comparer

Dès l'instant où plusieurs stimuli sont impliqués dans une comparaison, leur nombre peut influencer la manière dont ils seront traités et intégrés dans la généralisation ultérieure. Augier et Thibaut (2013) variaient le nombre d'items à comparer (items d'apprentissages)

Section 1: A theoretical review

présentés à des enfants de quatre et six ans. Les expérimentateurs ont confronté des situations de comparaison à deux ou quatre items d'apprentissages. Alors que les enfants de quatre ans généralisent le nouveau mot de la même façon avec deux ou quatre items d'apprentissage, ceux de six ans ont obtenus de meilleures performances de généralisation avec quatre items d'apprentissage.

En effet, l'augmentation du nombre d'items d'apprentissage augmente certes le nombre d'informations convergentes mais augmente aussi le coût cognitif d'une tâche (Zelazo & Frye, 1998; Andrews & Halford, 2002; Augier & Thibaut, 2013) notamment par le nombre de comparaisons à effectuer et à intégrer. Des différences exécutives pourraient expliquer les différences entre les deux groupes d'âge. Si, chez les enfants plus âgés, la performance augmente avec le nombre d'informations convergentes, chez les enfants de 4 ans, l'augmentation du nombre d'objets comparés augmentent le nombre de comparaisons nécessaires à une intégration de toute l'information disponible. Ce résultat a été répliqué par Thibaut et Witt (2015) dans le cas de noms de relation. La meilleure performance a été obtenue pour trois exemples d'apprentissage, supérieure à celles avec quatre items d'apprentissage.

La saillance des propriétés des objets de la tâche

Le même raisonnement peut s'appliquer à la saillance des propriétés. Lagarrigue et al. (en préparation) ont manipulé la saillance de la texture et de la forme d'objets inconnus. Dans une condition, forme et texture étaient saillantes, dans l'autre moins saillantes. Les performances de généralisation étaient inférieures (moins de choix fondés sur la dimension pertinente) dans la condition où la propriété pertinente était moins saillante. En outre, dans la condition moins saillante les différences entre les conditions avec ou sans comparaison étaient moins prononcées, notamment chez les plus jeunes. Par exemple, la situation de comparaison a donné des résultats proches du hasard et la différence entre comparaison inter et intra était plus faible. La recherche d'une dimension moins saillante, moins distinctive, nécessiterait une exploration plus systématique des stimuli, plus difficile pour les enfants de quatre ans.

La distance sémantique entre les items de la tâche

Un autre paramètre étudié pour mieux comprendre les tâches de comparaison-généralisation est la distance sémantique entre les items, une notion qui s'applique aux trois types de mots étudiés ici mais qui a surtout fait l'objet de recherches dans le cas des noms. Cette distance peut se calculer en termes du nombre de dimensions qui distinguent deux

Section 1: A theoretical review

objets, événements ou propriétés. Plus ce nombre est élevé plus les domaines conceptuels sont éloignés.

Dans le cas des noms, plusieurs études ont considéré l'effet de la distance sémantique entre les items. Thibaut et Witt (2017; voir également Thibaut & Witt, 2015, pour les termes relationnels ou Stansbury et al., 2019), ont proposé une tâche de comparaison pour l'apprentissage de catégories taxonomiques à des enfants de quatre et six ans, avec deux items d'apprentissage et trois items de transfert où l'un était un objet de même catégorie taxonomique, un autre était un objet perceptivement similaire des items d'apprentissage et le dernier un objet lié aux items d'apprentissage par une relation thématique. Les auteurs ont manipulé la distance sémantique entre les items d'apprentissage en présentant soit des paires d'items taxonomiquement proches (e.g., une pomme rouge et une pomme verte) soit des paires plus éloignées (e.g., une pomme rouge et une cerise). Ils ont également manipulé la distance entre les items d'apprentissage et l'item cible taxonomiquement relié (par exemple, une cible proche comme une banane ou une cible éloignée comme de la viande). D'abord, les performances de généralisation étaient meilleures pour des items d'apprentissage éloignés, ensuite, une interaction entre l'âge et la distance de généralisation a révélé une meilleure généralisation taxonomique dans la condition de généralisation éloignée et des items d'apprentissages éloignés chez les plus âgés, différence absente chez les plus jeunes. Comme le soulignent Augier et Thibaut (2013), la généralisation à un item éloigné est cognitivement plus coûteuse et pour cette raison moins accessible aux enfants de quatre ans. Une autre interprétation, non exclusive, serait que la généralisation éloignée repose sur une représentation plus inclusive, plus abstraite (dans le sens de plus détachée des items d'apprentissage) que les enfants de quatre ans n'ont pas encore construite.

Dans le cas des verbes, la distance entre les items de la tâche est associée à la variabilité de la tâche. Les travaux de Childers (2011; 2017) montrent que les enfants sont très sensibles aux variations de la tâche et comment les stratégies de résolution de la tâche sont affectées par ces variations (Childers, 2020). Childers (2011) compare les résultats de participants comparant un événement de référence soit à des événements où l'action change (balle cachée car poussée, balle cachée après un déplacement sur un plan incliné, balle cachée après qu'on l'a posée), soit à des événements dans lesquels le résultat change (balle termine cachée, ou visible sur la table, ou sur un support), soit à des événements identiques. L'augmentation de la variabilité entre les événements à comparer (si on compare la situation « événement identique » avec les conditions avec changements d'action ou de résultat)

Section 1: A theoretical review

améliore les performances de généralisation d'enfants de trente mois. Dans d'autres situations, cependant, la variabilité des paramètres a l'effet inverse et diminue la performance. Childers (2017) montre à des enfants de deux ans et demi plusieurs événements d'apprentissage variant sur la complexité de l'action à comparer. Ces événements sont soit effectués par le même acteur, soit effectués par trois personnes différentes. Quand l'action est dite simple (se taper la tête avec sa main), faire comparer des événements avec des acteurs différents améliore les performances de généralisation par rapports aux performances obtenues avec le même acteur. Cependant si l'action est dite complexe, au sens où elle implique des objets et pas seulement des parties du corps (taper un objet avec un marteau), alors faire varier l'acteur diminue les performances. Cette contradiction apparente est le signe que les interactions entre les paramètres des tâches et leurs effets sur les performances ou les stratégies des sujets sont subtiles et qu'il y a surtout une interaction entre variabilité (distance) et complexité. Il faut que les tâches soient le plus diversifiées possible pour permettre l'élaboration de la représentation la plus riche et souple possible sans pour autant que les situations de comparaison soient trop complexes pour les enfants.

Organisation de la tâche

Les tâches de comparaison de stimuli sont cognitivement complexes. Elles demandent une organisation temporelle des comparaisons, reposent sur un ensemble de comparaisons entre des stimuli, ou entre des actions réalisées sur des stimuli, mais aussi entre les items de généralisation et ceux d'apprentissage. L'étude des mouvements oculaires peut nous renseigner sur l'organisation de ces tâches et sur le rôle que l'enfant attribue aux différents stimuli. Par exemple, dans l'étude de Childers (2020), les enfants regardent plus longtemps l'outil quand celui-ci change d'un événement à l'autre mais se détournent de cet outil si celui-ci reste identique d'une réalisation de l'action à l'autre. En outre, dans la condition changement d'outil, leurs performances sont corrélées avec la longueur des temps de regard sur l'outil. La variabilité semble donc attirer l'attention des sujets sur les éléments pertinents pour la généralisation et influencer leurs stratégies de résolution et leurs performances.

Stansbury et al. (2019) ont étudié le mouvement des yeux dans une étude portant sur la généralisation des noms dans laquelle la distance sémantique entre items d'apprentissage variait. Les auteurs segmentaient les essais en trois tranches de temps (début, milieu, fin) afin d'appréhender le décours temporel de l'attention portée aux différents stimuli. Ils ont testé deux hypothèses d'organisation possible de la recherche d'information. La première prédisait que les enfants regardent d'abord les stimuli d'apprentissage puis comparent les items de

Section 1: A theoretical review

transfert entre eux pour décider lequel est la solution conceptuelle. La seconde prédisait que les enfants comparent d'abord les items d'apprentissage entre eux, puis chaque item de transfert séparément à ces items d'apprentissage, sans que les options proposées soient comparées entre elles. Les résultats montrent que les sujets suivent cette dernière organisation avec la recherche d'une relation entre les items d'apprentissage, confrontée ensuite à chaque option, une stratégie que les auteurs appellent une stratégie d'alignement, la première étant dénommée stratégie de projection d'une solution (Gentner & Colhoun, 2010). Contrairement aux hypothèses des auteurs, la distance entre items d'apprentissage et items test n'a pas influencé les stratégies de résolution des sujets. En revanche, le profil de recherche associé aux erreurs se démarquait de celui des bonnes réponses par une exploration moins systématique des stimuli.

L'organisation temporelle de présentation des items et les stratégies de résolution des sujets

On a également étudié l'effet de l'ordre temporel de présentation des items d'apprentissage (Christie & Gentner, 2010 ; Lawson, 2017; Son et al., 2011; Vlach et al., 2012), présentation simultanée des items d'apprentissage ou présentation séquentielle, avec un plus grand nombre de généralisations conceptuelles pour la présentation simultanée. C'est la comparaison directe des stimuli d'apprentissage présentés simultanément qui permet de diriger l'attention vers des similitudes et pas seulement la présentation de plusieurs exemples en tant que telle. Les contraintes mnésiques associées à une présentation séquentielle peuvent expliquer cette différence. En effet, les participants qui ont encodé le premier stimulus sur la base de dimensions saillantes ne pourront le comparer au second qu'en termes de ces dimensions saillantes. La présentation simultanée des stimuli d'apprentissage permet de constater que les dimensions saillantes ne sont pas pertinentes et peut contribuer à la recherche de dimensions moins saillantes, d'abord ignorées lors du premier encodage.

La plupart des auteurs décrivent le paradigme de comparaison sous la forme d'une étape d'apprentissage, suivie par une étape de généralisation. Cependant, on peut se demander si l'apprentissage et la généralisation ne sont pas plus entrelacées dans le sens où les enfants modifieraient leur conceptualisation durant la généralisation. Stansbury et al. (2020) manipulent le format de présentation des items de généralisation, séquentielle ou simultanée, et montrent que la présentation simultanée donne des performances supérieures à la présentation séquentielle chez des enfants de quatre ans. Les choix opérés par les enfants semblent suggérer qu'ils comparent les items de généralisation aux items d'apprentissage et

Section 1: A theoretical review

que ces comparaisons ultérieures conduisent à modifier leur représentation durant la généralisation. Cette interprétation dynamique de la généralisation est confirmée par l'étude de Stansbury et al. (2019) qui, par l'analyse des mouvements oculaires, révèle la présence de comparaisons entre items d'apprentissage et items de généralisation.

Ces résultats conduisent à des hypothèses sur le décours de la construction de la représentation mentale du concept cible. Soit, la représentation est extraite de la comparaison des seuls items d'apprentissage et est le seul support de la généralisation ultérieure : les participants généralisent à un objet qui a les mêmes caractéristiques que celles extraites à partir des items d'apprentissage. Soit, les sujets élaborent leur représentation avec les informations extraites de l'apprentissage mais aussi celles issues des comparaisons avec les items de généralisation. Ils construisent leur représentation progressivement, après avoir étudié l'ensemble des éléments de la tâche et de façon à créer une représentation qui intègre les items d'apprentissage et les informations venues des options proposées. Cette dernière hypothèse semble compatible avec les résultats précédents et présente une vision de la construction dynamique de la représentation qui s'adapterait aux items de généralisation.

Conclusion sur les mécanismes et les stratégies

L'étude des paramètres influençant la comparaison-généralisation des stimuli dans l'apprentissage de nouveaux mots permet de traduire ces tâches en mécanismes cognitifs et en stratégies de résolution. Les interactions observées entre l'âge et les paramètres de la tâche peuvent signifier l'implication des fonctions exécutives notamment l'inhibition et la flexibilité cognitive.

En effet, la comparaison dans les tâches d'apprentissage lexical nécessite une recherche des propriétés profondes (par opposition à superficielle) visant à extraire les informations pertinentes pour la généralisation comme l'indique les différences entre les profils de mouvements oculaires pour les bonnes et des mauvaises réponses. Augmenter la distance sémantique entre les items d'apprentissage semble favoriser la généralisation conceptuelle. Une distance plus importante permettrait de construire un espace sémantique plus large là où des items d'apprentissage sémantiquement proches limiteraient cet espace. En revanche, l'augmentation de la distance peut rendre l'appréhension des traits pertinent plus complexe notamment pour les enfants plus jeunes. Il y a donc un optimum de variabilité sémantique à l'apprentissage pour chaque groupe d'âge qui dépend du nombre d'items comparés, du nombre de comparaisons nécessaires à la saisie des traits pertinents, à leur

saillance. Cet optimum, selon nous, se définit largement en termes des compétences exécutives des enfants ciblés par la tâche.

Synthèse et conclusion

Dans cette revue de question, nous avons montré que les situations de comparaison par rapport aux situations de non-comparaison facilitent l'apprentissage des trois classes lexicales considérées, les noms, les verbes et les adjectifs. La contribution des deux types de comparaison analysés, intra et inter, varie selon le type de mot à apprendre. Si les situations de comparaison intra-catégorielle sont efficaces pour les trois classes de mots, les situations de comparaison inter-catégorie ont surtout montré des effets positifs pour l'apprentissage de nouveaux adjectifs et de nouveaux verbes alors qu'ils se sont révélés plus limités pour l'apprentissage de noms, du moins dans les situations d'apprentissage considérées ici.

Les résultats montrent également que l'efficacité de la comparaison dépend de la similarité sémantique des stimuli d'apprentissage qui contribue à la difficulté de la tâche. Les résultats indiquent que des entités d'apprentissage peu semblables sémantiquement rendent plus difficile la détection des dimensions pertinentes à certains âges mais contribuent aussi à une généralisation plus large. Enfin, les analyses des mouvements oculaires indiquent qu'une comparaison efficace consiste d'abord à rechercher les traits conceptuels qui unifient les items d'apprentissage avant de les comparer à chaque option proposée pour la généralisation. En effet, les erreurs se caractérisent, notamment, par une exploration moins systématique des items d'apprentissage mais aussi des items test. Les interactions entre l'âge et les paramètres de la tâche montrent que la complexité optimale de la comparaison dépend aussi des caractéristiques cognitives des enfants: les gains de la comparaison peuvent diminuer ou disparaître lorsque la complexité sémantique et procédurale de la tâche outrepassent les caractéristiques cognitives des enfants d'un âge donné (Augier & Thibaut, 2013; Thibaut & Witt, 2015)

Limites

Bien que certaines études sur l'apprentissage d'un nouveau nom aient été réalisées avec du matériel non-familier, la plupart des recherches ont été faites avec des items familiers des enfants. Ces études testaient si les enfants étaient capables de généraliser les items d'apprentissage sur la base d'une relation taxonomique plutôt que perceptive. Ce paradigme mobilise des connaissances déjà disponibles et une interprétation de la tâche serait que les enfants associent une étiquette connue (par exemple pomme) aux objets d'apprentissage (deux

Section 1: A theoretical review

pommes) en utilisant les concepts déjà connus correspondant, une éventualité d'ailleurs discutée (et rejetée) par Namy et Gentner (2002). Dans cette perspective, la situation de comparaison ne crée pas de nouvelles dimensions mais donne une pondération différente à une dimension qui reste moins saillante dans la situation de non-comparaison. A l'inverse, une approche fondée sur des éléments non familiers, c'est-à-dire d'éléments pour lesquels l'enfant n'a aucune connaissance à priori, semble intéressante pour caractériser la construction de nouvelles dimensions conceptuelles. Les deux types d'études, celles utilisant des stimuli familiers et celles reposant sur des stimuli non familiers, montrent le même avantage de la comparaison. Cependant, celles avec du matériel non familier restent trop peu nombreuses, limitées à des dimensions perceptives simples et, souvent, à des catégories de niveau de base. Elles n'ont pas encore été utilisées pour étudier les effets de la comparaison dans l'apprentissage d'adjectifs ou de verbe.

Les analyses portant sur la contribution respective de l'intra et de l'inter (contraste) proposées par Hammer et al. (2008, 2009) ont été illustrées par des situations expérimentales d'apprentissage très contrôlées et, souvent, avec des stimuli artificiels créés à partir d'un nombre limité de dimensions (par exemple la forme, la couleur et la texture). Lorsque l'on présente un objet de contraste qui montre la non-pertinence d'un ou plusieurs critères (la forme et la couleur), cette information permet de déduire la valeur diagnostique de l'autre dimension (la texture) car l'espace des dimensions est limité. Cette opération devient impossible dans le réel où les stimuli peuvent être décrits en termes d'un nombre infini de dimensions (Murphy & Medin, 1985, pour une discussion). L'analyse du contraste reste à faire dans des situations hautement multidimensionnelles ou, mieux, l'analyse du rôle du contraste devrait se faire en manipulant le nombre de dimensions des stimuli. Enfin, si les études ont porté sur les noms d'objet et de relation, sur les verbes et des adjectifs nommant des propriétés perceptives d'objets, à notre connaissance, il n'y a pas d'étude systématique de mots appartenant à d'autres classes grammaticales.

Perspectives

Plusieurs axes de perspectives peuvent être envisagés dans le cadre de la comparaison. Premièrement, l'implication des fonctions exécutives n'a été testée, à ce jour, que par la manipulation des poids exécutifs des tâches. Une approche psychométrique classique, complémentaire, pourrait être envisagée. La manipulation des fonctions exécutives est également une option. Une possibilité serait de bloquer la mise à jour (updating) afin d'étudier son rôle, en augmentant les contraintes sur la mémoire de travail en utilisant une

Section 1: A theoretical review

double tâche, tel que cela est classiquement réalisé pour aborder cette question (Baddeley, 2007).

Deuxièmement, nous avons évoqué à plusieurs reprises l'asymétrie entre les conditions intra et inter dans le cas des noms. A l'heure actuelle, aucune étude n'a comparé directement les effets de ces deux conditions pour l'apprentissage des verbes ou des adjectifs. Si les deux types de comparaison sont efficaces pour ces deux classes de mots, on ne peut pas conclure sur la supériorité de l'une sur l'autre. Nous pourrions donc imaginer une étude qui répondrait à cette question et qui confirmerait, ou non, que les deux types de comparaison sont autant efficaces pour les verbes et les adjectifs.

Troisièmement, les études sur l'apprentissage des verbes reposent sur une présentation dynamique des stimuli car les différentes étapes des événements d'entraînement ne peuvent être présentées simultanément. Pour les noms, la présentation séquentielle des items d'apprentissage produit des résultats moins convaincants qu'une présentation simultanée. Par contre, dans le cas des noms relationnels, on pourrait comparer une présentation dynamique de la relation à une comparaison statique afin d'évaluer les mérites de chacune.

Quatrièmement, si cette revue a traité des noms, des verbes et des adjectifs, d'autres classes lexicales, telles que les pronoms, les conjonctions ou les prépositions n'ont pas été étudiées dans ce contexte par les chercheurs. Les prépositions spatiales, notamment, qui reposent essentiellement sur le contraste entre des dimensions précises, sont une cible idéale pour ces recherches.

Enfin, de manière générale, les recherches sur l'apprentissage du langage par la comparaison pourraient bénéficier des recherches équivalentes en dehors de ce domaine, que ce soit dans la littérature sur les apprentissages scolaires ou des apprentissages d'une expertise. De nombreux facteurs ont été étudiés comme le montre la méta-analyse d'Alfieri et al. (2013). Cette analyse suggère l'influence de facteurs contextuels, comme le délai ou l'existence d'une explication du contenu. Certains de ces facteurs pourraient faire l'objet d'une application dans le domaine langagier.

Bibliographie

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist*, 48(2), 87-113.

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology*, 45(2), 153–219.
[https://doi.org/10.1016/S0010-0285\(02\)00002-6](https://doi.org/10.1016/S0010-0285(02)00002-6)

Ankowski, A. A., Vlach, H. A., & Sandhofer, C. M. (2013). Comparison versus contrast: Task specifics affect category acquisition. *Infant and Child Development*, 22(1), 1–23.

Augier, L., & Thibaut, J. P. (2013). The benefits and costs of comparisons in a novel object categorization task: Interactions with development. *Psychonomic Bulletin & Review*, 20(6), 1126–1132. <https://doi.org/10.3758/s13423-013-0436>

Baddeley, A. (2007). *Working memory, thought, and action* (Vol. 45). OuP Oxford.

Bassano D. (2000/03), La constitution du lexique : le « développement lexical précoce ». In M. Kail & M. Fayol, *L'acquisition du langage*, vol. 1, *Le langage en émergence*, pp. 137-168, Paris : PUF

Bassano, D., Labrell, F., Champaud, C., Lemétayer, F., & Bonnet, P. (2005). Le dlpf: un nouvel outil pour l'évaluation du développement du langage de production en français. *Enfance*, 57(2), 171-208.

Behrend, D. A. (1995). Processes involved in the initial mapping of verb meanings. *Beyond names for things: Young children's acquisition of verbs*, 251-273.

Booth, A. E., & Waxman, S. R. (2003). Mapping words to the world in infancy: Infants' expectations for count nouns and adjectives. *Journal of Cognition and Development*, 4(3), 357–381. https://doi.org/10.1207/S15327647JCD0403_06

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, 75, 1115–1139.

Boucheix, J. M., Lowe, R. K., & Thibaut, J. P. (2020). A developmental perspective on young children's understandings of paired graphics conventions from an analogy task. *Frontiers in Psychology*.

Section 1: A theoretical review

Childers, J. B. (2008). *The structural alignment and comparison of events in verb acquisition*. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 30, No. 30).

Childers, J. B. (2011). Attention to multiple events helps two-and-a-half-year-olds extend new verbs. *First Language, 31*(1), 3–22. <https://doi.org/10.1177/01427237110361825>

Childers, J. B., & Paik, J. H. (2009). Korean- and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language, 36*(1), 201–224. <https://doi.org/10.1017/S0305000908008891>

Childers, J. B., & Tomasello, M. (2006). 12 Are Nouns Easier to Learn Than Verbs? Three Experimental Studies. *Action meets word: How children learn verbs*, 311.

Childers, J. B., Hirshkowitz, A., & Benavides, K. (2014). Attention to explicit and implicit contrast in verb learning. *Journal of cognition and development, 15*(2), 213-237.

Childers, J. B., Lalani, S., Porter, B., Arriazola, S., Tovar-Perez, P., & Cutilletta, B. (2019). Studying Children’s Verb Learning across Development. *Conducting Research in Developmental Psychology: A Topical Guide for Research Methods Utilized Across the Lifespan*, 104.

Childers, J. B., Paik, J. H., Flores, M., Lai, G., & Dolan, M. (2017). Does variability across events affect verb learning in English, Mandarin, and Korean?. *Cognitive science, 41*, 808-830.

Childers, J. B., Parrish, R., Olson, C. V., Burch, C., Fung, G., & McIntyre, K. (2016). Early verb learning: How do children learn how to compare events? *Journal of Cognition and Development, 17*(1), 41–66. <https://doi.org/10.1080/15248372.2015.1042580>

Childers, J. B., Porter, B., Dolan, M., Whitehead, C. B., & McIntyre, K. P. (2020). Does children’s visual attention to specific objects affect their verb learning? *First Language, 40*(1), 21–40. <https://doi.org/10.1177/0142723719875575>

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11*(3), 356–373. <https://doi.org/10.1080/15248371003700015>

Clark, E. V. (2009). *First language acquisition*. Cambridge, Cambridge University Press. <https://doi.org/10.1017/CBO9780511806698>

Section 1: A theoretical review

Clark, E.V. (1995). *The lexicon in acquisition*. Cambridge, Cambridge University Press.

Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development, 74*(1), 168–178. <https://doi.org/10.1111/1467-8624.00528>

Forbes, J. N., & Farrar, M. J. (1995). Learning to represent word meaning: What initial training events reveal about children's developing action verb concepts. *Cognitive Development, 10*(1), 1–20. [https://doi.org/10.1016/0885-2014\(95\)90016-0](https://doi.org/10.1016/0885-2014(95)90016-0)

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780195154061.001.0001>

Gelman, S. A. (2013). Concepts in development. In P. D. Zelazo (Ed.), *The Oxford handbook of developmental psychology*. New York: Oxford University Press.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S. A. Kuczaj (Ed.), *Language development: Language, thought and culture* (Vol.2, PP. 301-324) Hillsdale, NJ: Lawrence Erlbaum

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155-170

Gentner, D., & Colhoun, J. (2010). Analogical processes in human thinking and learning. In *Towards a theory of thinking* (pp. 35-48). Springer, Berlin, Heidelberg.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45-56

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487–513. [https://doi.org/10.1016/S0885-2014\(99\)00016-](https://doi.org/10.1016/S0885-2014(99)00016-)

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*(6), 297–301. <https://doi.org/10.1111/j.1467-8721.2006.00456>.

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82*(4), 1173–1188. <https://doi.org/10.1111/j.1467-8624.2011.01599>.

Section 1: A theoretical review

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*(3), 285–307.

<https://doi.org/10.1080/1524837070144643>

Graham, S. A., Cameron, C. L., & Welder, A. N. (2005). Preschoolers' extension of familiar adjectives. *Journal of Experimental Child Psychology, 91*(3), 205–226.

<https://doi.org/10.1016/j.jecp.2005.03.00>

Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107*(3), 280–290. <https://doi.org/10.1016/j.jecp.2010.04.01>

Gropen, J., Pinker, S., Hollander, M., & Goldberg, R. (1991). Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure. *Cognition, 41*(1-3), 153-195.

Hall, D. G., & Rhemtulla, M. (2014). Young children's use of contrast in word learning: the case of proper names. *Journal of Cognition and Development, 15*(4), 551-568.

Hall, D. G., Waxman, S. R., & Hurwitz, W. M. (1993). How two- and four-year-old children interpret adjectives and count nouns. *Child Development, 64*(6), 1651–1664.

<https://doi.org/10.2307/1131461>

Hammer, R., Bar-Hillel, A., Hertz, T., Weinshall, D., & Hochstein, S. (2008). Comparison processes in category learning: From theory to behavior. *Brain Research, 1225*, 102–118. <https://doi.org/10.1016/j.brainres.2008.04.079>

Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference? *Cognition, 112*(1), 105–119. <https://doi.org/10.1016/j.cognition.2009.03.012>

Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extension. *Child Development, 82*(2), 674–686.

<https://doi.org/10.1111/j.1467-8624.2010.01567.x>

Hirsh-Pasek, K., Golinkoff, R. M., & Hollich, G. (2000). An emergentist coalition model for word learning. *Becoming a word learner: A debate on lexical acquisition*, 136-164.

Hoff, E. (2013). *Language development*. Wadsworth, Cengage Learning.

Section 1: A theoretical review

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, 8(2), 113–139. [https://doi.org/10.1016/0885-2014\(93\)90008-S](https://doi.org/10.1016/0885-2014(93)90008-S)

Kaminski, J. A., & Sloutsky, V. M. (2013). Extraneous perceptual information interferes with children's acquisition of mathematical knowledge. *Journal of Educational Psychology*, 105(2), 351–363. <https://doi.org/10.1037/a0031040>

Klibanoff, R. S., & Waxman, S. R. (2000). Basic level object categories support the acquisition of novel adjectives: Evidence from preschool-aged children. *Child Development*, 71(3), 649–659. <https://doi.org/10.1111/1467-8624.00173>

Kok, E. M., de Bruin, A. B., Robben, S. G., & van Merriënboer, J. J. (2013). Learning radiological appearances of diseases: Does comparison help? *Learning and Instruction*, 23, 90–97. <https://doi.org/10.1016/j.learninstruc.2012.07.004>

Kotovsky, L., & Gentner, D. (1996). Comparison and Categorization in the Development of Relational Similarity. *Child Development*, 67(6), 2797-2822.

Kucker, S. C., Samuelson, L. K., Perry, L. K., Yoshida, H., Colunga, E., Lorenz, M. G., & Smith, L. B. (2019). Reproducibility and a unifying explanation: Lessons from the shape bias. *Infant Behavior and Development*, 54, 156–165. <https://doi.org/10.1016/j.infbeh.2018.09.011>

Lagarrigue Y., Augier L., & Thibaut J.P. (2020). Learning novel names with multiple examples: dimensional distinctiveness and comparison formats. Article en preparation.

Lagarrigue Y., Thibaut J.P. (2020). Apprendre et généraliser de nouveaux noms, les leçons de la comparaison : une revue de la littérature. *Enfance*. In press.

Landau, B. (1994). Object shape, object name and object kind: representation and development. D. Medin (Ed.) *Psychology of Learning and Motivation: Advances in Research and Theory* (Vol 31), 253-304.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3(3), 299–321. [https://doi.org/10.1016/0885-2014\(88\)90014-7](https://doi.org/10.1016/0885-2014(88)90014-7)

Lawson, C. A. (2017). The influence of task dynamics on inductive generalizations: How sequential and simultaneous presentation of evidence impacts the strength and scope of

Section 1: A theoretical review

property projections. *Journal of Cognition and Development*, 18(4), 493–513.

<https://doi.org/10.1080/15248372.2017.1339707>

Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). A Unified Theory of Word Learning: Putting Verb Acquisition in Context. *Action meets word: How children learn verbs*, 364.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, 11(4), 628-634. <https://doi.org/10.1111/j.1467-7687.2008.00707.x>

Mareschal, D., Quinn, P. C., & Lea, S. E. G. (Eds.). (2010). *Oxford series in developmental cognitive neuroscience. The making of human concepts*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199549221.001.0001>

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.

Markman, E. M. (1994). Constraints on word meaning in early language acquisition. *Lingua*, 92(1-4), 199–227. [https://doi.org/10.1016/0024-3841\(94\)90342-5](https://doi.org/10.1016/0024-3841(94)90342-5)

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual review of psychology*, 32(1), 89-115.

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, 84(3), 267–293. [https://doi.org/10.1016/S0010-0277\(02\)00047-1](https://doi.org/10.1016/S0010-0277(02)00047-1).

Murphy, G. L. (2002). *The big book of concepts*. MIT Press.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological review*, 92(3), 289.

Namy, L. L., & Clepper, L. E. (2010). The differing roles of comparison and contrast in children's categorization. *Journal of Experimental Child Psychology*, 107(3), 291–305. <https://doi.org/10.1016/j.jecp.2010.05.013>

Section 1: A theoretical review

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology. General*, *131*(1), 5–15. <https://doi.org/10.1037/0096-3445.131.1.5>

Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognitie, Creier, Comportament/Cognition. Brain and Behavior*, *11*(4), 647–659.

Nelson, K., & Gruendel, J. M. (1981). Generalized event representations: Basic building blocks of cognitive development. In M. E. Lamb & A. L. Brown (Eds.), *Advances in developmental psychology* (Vol. 1, pp. 131–158). Erlbaum.

Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of experimental psychology: General*, *115*(1), 39.

Quine, W. V. (1960). Carnap and logical truth. *Synthese*, *12*(4), 350–374. <https://doi.org/10.1007/BF00485423>

Rosch, E. (1978). Principles of classification. *Cognition and categorization*, 27–48.

Scott, R. M., & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes*, *24*, 777–803. [doi:10.1080=01690960802573236](https://doi.org/10.1080=01690960802573236)

Scott, R. M., & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, *122*(2), 163–180. <https://doi.org/10.1016/j.cognition.2011.10.010>

Sloutsky, V. (2015). Conceptual development. In L. S. Liben, U. Müller, & R. M. Lerner (Eds.), *Handbook of child psychology and developmental science: Cognitive processes* (p. 469–518). John Wiley & Sons Inc

Smith, L. B. (1989). A model of perceptual classification in children and adults. *Psychological Review*, *96*(1), 125–144. <https://doi.org/10.1037/0033-295X.96.1.125>

Snape, S., & Krott, A. (2018). The benefit of simultaneously encountered exemplars and of exemplar variability to verb learning. *Journal of Child Language*, *45*(6), 1412–1422. <https://doi.org/10.1017/S0305000918000119>

Section 1: A theoretical review

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2011). Connecting instances to promote children's relational reasoning. *Journal of Experimental Child Psychology*, *108*(2), 260–277. <https://doi.org/10.1016/j.jecp.2010.08.011>

Stansbury, E., Witt, A., & Thibaut, J. P. (2019). *Children's Generalization of Novel Object Names in Comparison Contexts: An eye tracking analysis*. In A.K. Goel, C.M. Seifert, & C. Freksa (Eds.), *In Proceedings of the 41st Annual Meeting of the Cognitive Science Society*, (pp. 2871-2877). Cognitive Science Society.

Stansbury, E., Witt, A., & Thibaut, J. P. (2020). Generalization of Novel Object Names in Comparison Contexts in a yes-no paradigm by young children. When the rate of stimulus presentation matters. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 2288 - 2294). Cognitive Science Society.

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian and Mandarin. *Journal of Child Language*, *24*, 535–565.

Thibaut, J. P. (1997). Similarité et catégorisation. *L'année psychologique*, *97*(4), 701-736.

Thibaut, J. P., & Witt, A. (2015). Young children's learning of relational categories: Multiple comparisons and their cognitive constraints. *Frontiers in Psychology*, *6*, 643. <https://doi.org/10.3389/fpsyg.2015.00643>

Thibaut, J. P., & Witt, A. (2017). *Generalizing novel names in comparison settings: Role of conceptual distance during learning and at test*. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society* (pp. 3314-3319). Austin, TX: Cognitive Science Society.

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511527678>

Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(1), 246–254. <https://doi.org/10.1037/a0025260>

Section 1: A theoretical review

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, *36*(5), 571–581. <https://doi.org/10.1037/0012-1649.36.5.571>

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty-four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive psychology*, *59*(1), 67-95.

Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science*, *7*(4), 121–126. <https://doi.org/10.1111/1467-8721.ep10774761>

**SECTION 2: INVESTIGATING WORD
GENERALIZATION STRATEGIES.**

Introduction

In novel word learning and generalization tasks, mappings are made between the task items to explore them and align their common features. Other mappings are made between the task's items and existing concepts: as each item is processed and recognized, its features are mapped to the concept's corresponding features. In the following section we will present empirical work which aimed at getting a better understanding of novel word generalization strategies and the processes that may support them, in other words of the mappings and processes involved in concept learning and use. The first two studies – presented in Chapters 1 and 2 - explore generalization strategies thanks to eye tracking data, while processes that support generalization strategies were explored in an executive function framework thanks to cognitive assessments that are presented in Chapter 3.

Chapters 1 and 2 stem from the observation that most research on novel word generalization has solely considered and analyzed children's generalization choice, as a function of the settings children were in. Consequently, all knowledge of children's generalization strategies comes from inferences about strategies, attentional course and the process that may be at play from the observed choices rather than data revealing the strategies directly. However, the challenge in word generalization tasks is undoubtedly managing to attend to the conceptual information rather than other sources of information, which makes understanding children processes in such tasks, where they look, the mappings they make, key investigation topic. To reveal the strategies, we used eye tracking data.

Eye movements provide a precise reflection of the interactions between cognitive processes and the visual external stimuli (Kahneman, 2011; Rayner et al., 2018). Eye position and attention are strongly correlated (Fischer & Breitmeyer, 1987; Klein et al., 1992; Remington, 1980; Shepherd et al., 1986). Attentional movements and eye saccades are also considered coupled (Deubel & Schneider, 1996).

The major asset of eye tracking data is that it provides a real-time view of what participants attend to while solving a task. For this reason eye movements have been used to investigate how children are learning across examples in the case of verbs (Childers et al., 2020); the time course of analogical reasoning in adults (Glady et al., 2016; Thibaut et al., 2022); reasoning strategies in propositional analogy problems (Vendetti et al., 2017); relational and object matching processes in picture analogy solving tasks (Gordon & Moser, 2007); solving strategies in matrix solving tasks (Chen et al., 2016); geometric analogy task solving strategies and individual differences (Bethell-Fox et al., 1984); individual differences in children's analogical reasoning strategies (Starr et al., 2018), and the development of analogical reasoning (Thibaut & French, 2016).

Eye movement indexes are needed to investigate novel noun generalization strategies. Studies of problem solving have found that looking time to critical items provides the most informative measures of underlying cognitive processes (Hodgson et al., 2000; Knoblich et al., 2001) while it does not always appear a significant indicator. In relational and object-matching processes, Gordon and Moser (2007) found that saccade patterns between objects were very sensitive while time spent looking at objects wasn't. Even though looking times and saccade patterns sensitivity isn't systematic in prior research, both have participated in revealing participants' problem solving or matching strategies. We will therefore consider looking times and saccades between items in our empirical work to reveal participants' word learning and generalization strategies temporal dynamics.

Our suggestion is that the type of strategies used by participants and the organization of these strategies' steps will add to our understanding of how participants come to a generalization decision. In other words, we suggest that the data obtained will be a first insight into what participants are doing when learning and using concepts to generalize a word. Differences - if any exist - between participants' strategies or between strategies that led to different outcomes will also be a valuable source of information about what drives correct taxonomic generalization.

Novel noun generalization strategies will be investigated for object names (i.e., nouns) (Chapter 1) and relational nouns (Chapter 2). We can note straight away that the difference between these two tasks lies not in the cognitive task (i.e., learning a noun, extracting the concept, generalizing the word by inferences based on the concept) but in the nature of the concept that is being learnt and used during the task. In relational noun generalization, participants work with a concept that describes the relation between items and is consequently not linked to the items' features. In the case of noun generalization, the concepts targeted by the task are taxonomic concepts. These concepts are feature based, even though the lexicalized features such as the name are symbolic lexicalized information, other features are intrinsic to the items seen in the task. The difference between the relational noun generalization task and the noun generalization task could be summarized as a concept level of abstraction difference.

Both these studies investigate generalization strategies in multiple age groups in order to reveal any developmental changes in the strategies themselves or in their execution. In the case of novel noun generalization, we investigated generalization in 5-, 6-, and 8- year-old children. In the case of novel relational nouns generalization, we compared 5-year-old children's strategies with those of adults. The rationale behind the choice of age groups was to investigate strategies in young children - 5-years old being the youngest age for which eye tracking data was reliable in the present tasks - and compare to mature strategies to witness developmental change if one exists.

Our second investigation was to study processes that underlie novel noun and novel relational noun generalization tasks (Chapter 3). Our suggestion is that comparing multiple learning items and choosing a generalization match, while neglecting irrelevant dimensions including salient dimensions such as perceptual similarities involves flexibility, working memory, inhibition, and decision making. We investigate individual differences in participants' generalization scores and strategies as a function of their levels of executive development assessed by executive function tests.

References

- Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984). Adaptive reasoning : Componential and eye movement analysis of geometric analogy performance. *Intelligence*, 8(3), 205-238. [https://doi.org/10.1016/0160-2896\(84\)90009-6](https://doi.org/10.1016/0160-2896(84)90009-6)
- Chen, Z., Honomichl, R., Kennedy, D., & Tan, E. (2016). Aiming to Complete the Matrix : Eye-Movement Analysis of Processing Strategies in Children's Relational Thinking. *Developmental Psychology*, 52. <https://doi.org/10.1037/dev0000113>
- Childers, J. B., Porter, B., Dolan, M., Whitehead, C. B., & McIntyre, K. P. (2020). Does children's visual attention to specific objects affect their verb learning? *First Language*, 40(1), 21-40. <https://doi.org/10.1177/0142723719875575>
- Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition : Evidence for a common attentional mechanism. *Vision Research*, 36(12), 1827-1837. [https://doi.org/10.1016/0042-6989\(95\)00294-4](https://doi.org/10.1016/0042-6989(95)00294-4)
- Fischer, B., & Breitmeyer, B. (1987). Mechanisms of visual attention revealed by saccadic eye movements. *Neuropsychologia*, 25(1, Part 1), 73-83. [https://doi.org/10.1016/0028-3932\(87\)90044-3](https://doi.org/10.1016/0028-3932(87)90044-3)
- Glady, Y., French, R. M., & Thibaut, J.-P. (2016). *Comparing competing views of analogy making using eye-tracking technology*.
- Gordon, P. C., & Moser, S. (2007). Insight into analogies : Evidence from eye movements. *Visual Cognition*, 15(1), 20-35. <https://doi.org/10.1080/13506280600871891>
- Hodgson, T. L., Bajwa, A., Owen, A. M., & Kennard, C. (2000). The Strategic Control of Gaze Direction in the Tower of London Task. *Journal of Cognitive Neuroscience*, 12(5), 894-907. <https://doi.org/10.1162/089892900562499>
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Klein, R., Kingstone, A., & Pontefract, A. (1992). Orienting of Visual Attention. In K. Rayner (Éd.), *Eye Movements and Visual Cognition : Scene Perception and Reading* (p. 46-65). Springer. https://doi.org/10.1007/978-1-4612-2852-3_4
- Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition*, 29(7), 1000-1009. <https://doi.org/10.3758/BF03195762>
- Rayner, K., Shen, D., Bai, X., & Yan, G. (2018). *Cognitive and cultural influences on eye movements*. CRC Press.

- Remington, R. W. (1980). Attention and saccadic eye movements. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 726-744.
<https://doi.org/10.1037/0096-1523.6.4.726>
- Shepherd, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology Section A*, 38(3), 475-491. <https://doi.org/10.1080/14640748608401609>
- Starr, A., Vendetti, M. S., & Bunge, S. A. (2018). Eye movements provide insight into individual differences in children's analogical reasoning strategies. *Acta Psychologica*, 186, 18-26. <https://doi.org/10.1016/j.actpsy.2018.04.002>
- Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions : A developmental investigation with eye-tracking. *Cognitive Development*, 38, 10-26.
- Thibaut, J.-P., Glady, Y., & French, R. M. (2022). Understanding the What and When of Analogical Reasoning Across Analogy Formats : An Eye-Tracking and Machine Learning Approach. *Cognitive Science*, 46(11), e13208.
<https://doi.org/10.1111/cogs.13208>
- Vendetti, M. S., Starr, A., Johnson, E. L., Modavi, K., & Bunge, S. A. (2017). Eye Movements Reveal Optimal Strategies for Analogical Reasoning. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00932>

Chapter 1 - How children generalize novel nouns: an eye-tracking analysis of their generalization strategies.

The following chapter contains the article: E. Stansbury, A. Witt, P. Bard, J-P. Thibaut, (submitted), How children generalize novel nouns: an eye tracking analysis of their generalization strategies. *PlosOne*

Short title: Noun generalization strategies

How children generalize novel nouns:
An eye-tracking analysis of their generalization strategies

Eleanor Stansbury^{1*}, Arnaud Witt¹, Patrick Bard¹, Jean-Pierre Thibaut¹

¹Laboratoire d'Étude de l'Apprentissage et du Développement, CNRS UMR 5022,
Université de Bourgogne, Dijon, France

*Corresponding Authors

E-mail : Eleanor.Stansbury@u-bourgogne.fr and jean-pierre.thibaut@u-bourgogne.fr

Abstract

Recent research has shown that comparisons of multiple stimuli which are associated with the same novel noun favors taxonomic generalization of this noun. These findings contrast with single-stimulus learning, during which children adopt so-called lexical biases. However, little is known about the underlying search strategies. The present experiment provides an eye-tracking analysis of novel word learning by comparison. We manipulated both the conceptual distance between the two learning items, i.e., children saw examples which they could use to learn the word (e.g., learning items; two bracelets versus a bracelet and a watch), and the conceptual distance between the learning items and the taxonomically related items among the available generalization options (e.g., the correct taxonomic generalization answer; a pendant versus a bow tie). We tested 5-, 6- and 8-year-old children's taxonomic (versus perceptual and thematic) generalization of novel names for objects. The search patterns showed that participants first focused on the learning items and then compared them with each of the possible choices and spent far less time comparing the various options with one another and this search profile remained stable across age groups. Data also revealed that early comparisons, (i.e., reflecting alignment strategies) predicted generalization performance. We discuss four search strategies as well as the effect of age and conceptual distance on these strategies.

Key Words: comparison, novel noun extension, generalization, eye-tracking, nouns, concepts, conceptual distance, search strategies

Introduction

One remarkable feature of novel word learning is that it is achieved with a small number of learning stimuli and a small number of associations between the stimulus and the corresponding noun [1,2,3]. Understanding how children make sense of this paucity of information during novel word learning and which situations favor accurate word generalization is crucial for understanding language development. This problem has received considerable attention in recent decades and there are various competing theories regarding the explanatory factors [4,5,6,7,8,9,10,11].

One prominent view of word learning and generalization is that a set of lexical constraints bias children towards particular dimensions of the stimuli [12, 13, 14]. Several lexical biases that minimize the number and type of dimensions children use as a generalization basis have been described in the literature [15, 16, 17]. For example, the shape bias leads children to generalize novel nouns towards targets with the same shape rather than the same color, size or texture [18, 19]. Given the large number of dimensions along which an object can *a priori* be described, constraining the number of potential dimensions novel words might refer to may help improve children's conceptually-based generalization.

Most authors would acknowledge the existence of the shape bias even though they may explain it in different ways, considering it as a bottom-up process [20] or as being susceptible to early conceptual influences (see [21], for an extensive discussion of the literature). Smith and colleagues [20] have proposed a four-step bottom-up model of how children learn to relate object names and attention to shape. At the end of this associative learning process, the object's shape becomes systematically associated with the object word. In Step 1, children associate names with individual objects (e.g., "this one is a ball"). In Step 2, first-order generalizations about individual object categories are made as children associate each category with a shape (e.g., balls are round). In Step 3, a higher-order generalization is made across learned categories and relates to the structure of object categories, which are

labelled with a name—that is, these categories are defined by shape similarities. In Step 4, children have thus already learned to attend to shape when they learn novel object names and can rapidly learn novel names.

Even though various theories exist as to what makes a dimension relevant [22, 23, 24, 25, 26, 27, 28, 29], a topic which is beyond the scope of the present project, the most important aspect of the process underpinning children’s word generalization is that they come to pay attention to specific dimensions unifying the category (here, object categories). In this respect, to gain an insight into the processes that support novel noun generalization, it is necessary to investigate how children’s attention changes during a novel noun generalization task.

Here, we concentrate on one novel noun learning situation that involves comparing several stimuli, in which lexical biases might be less appropriate or might guide the child towards irrelevant dimensions. For example, the shape bias would be irrelevant when shape is not relevant as a unifying cue [21, 30, 31].

Novel noun Generalization tasks

Novel noun generalization tasks are commonly used to investigate the way children extend novel nouns. In such tasks, children are usually shown a learning (otherwise known as a standard or reference) item. Past studies have made use of both pictures of objects and replications of objects. The item is associated with (often) a nonword (e.g. “*This is a buxi*” while showing a bracelet). The children are asked to choose another item among a set of options that could also have the same noun (e.g., “Which one is also a *buxi*?” among, for example, a necklace and a tyre in Fig 1.A, below). The specific options from which the children must choose depend on what the researcher wishes to study: one is a “correct” item that is taxonomically related to the learning stimuli (e.g. another piece of jewelry in the

present example), while the others are distractors that can be semantically or perceptually related to the learning stimuli [19, 32, 33, 34, 35, 36, 37].

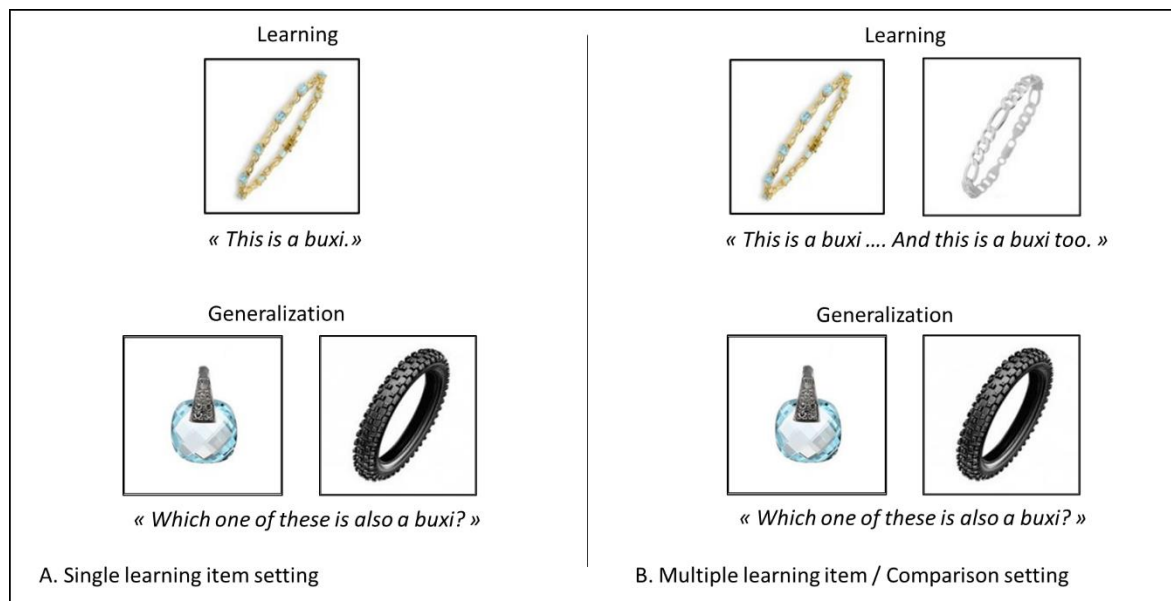


Fig 1. Novel noun generalization tasks in A. a single learning item setting and B. a multiple learning item (comparison design).

It is important to stress that much of the evidence regarding novel noun learning and generalization has been obtained with a single-object task. Recent studies have focused on variations of this learning task that might constrain children’s learning of word meaning. In this respect, it has been shown that the possibility of comparing multiple learning items (e.g., two or more objects belonging to the same category such as two apples or two pieces of fruit) associated with a single noun, rather than a single learning item, favors taxonomic responding or the use of less salient unifying dimensions [32] (see Fig 1.B). To illustrate, in their seminal experiment, Gentner and colleagues created a no-comparison and a comparison condition. In the no-comparison condition, they followed the above design and presented 4-year-olds with one familiar object accompanied by a novel (made-up) name (e.g. “*buxi*” in Fig. 1). They then showed the children two pictures, one of which was a taxonomic match and the other a perceptual match, and asked them which one had the “same name as the learning stimuli”

(e.g. “*Which one is also a buxi?*”). The comparison condition was identical except that there were two learning stimuli, which were both presented with the same name. Results showed that most children selected the taxonomic match significantly more often in the comparison than in the no-comparison condition, in which the majority of children selected the perceptual match.

A large body of research has demonstrated the benefits of comparison designs for learning novel object names [38, 39, 40], adjectives [41], action verbs [42], relational nouns [43, 44] (see [45] for a meta-analysis; [46] for a synthesis). Considerable attention has also been paid to the conditions under which comparisons lead to better learning and generalization. For example, the effect of the number of items compared [37, 44], the perceptual similarity between the learning stimuli [47], or the conceptual distance between them [48, 49] have all been extensively studied.

According to Gentner and colleagues, comparison conditions in novel-noun learning tasks elicit a deeper encoding of the compared learning items and lead children to progressively ignore irrelevant perceptual dimensions by emphasizing non-obvious conceptual properties [50, 51]. It is thought that this type of encoding is supported by processes that align features between learning items. Alignment is an “umbrella” term which is widely used in analogical reasoning to describe how participants compare a base domain and a target domain in order to find both perceptual and higher-order relational similarities between domains [50]. Authors consider that, in the novel noun learning by comparison task, participants also align stimuli as they try to find commonalities, either perceptual or less obvious, deeper, ones [52, 53, 54, 55, 56]. Comparisons of multiple learning items at the beginning of a novel noun generalization task promote the alignment of their common features [32] and labels [51]. A more specific hypothesis is that the comparison process would start with an alignment of the most salient perceptual features. These earlier alignments would

promote new comparisons that then lead to the discovery and alignment of less salient conceptual commonalities. Thus, the comparison process highlights conceptual features that would have remained unnoticed in a no-comparison situation; and the common representation that is extracted from this alignment process between item features is more conceptually based than the one that would have been built in a no-comparison learning condition. The fact that comparison situations lead to more conceptually-based (taxonomic in the case of nouns) choices when a noun is associated (rather than simply compared) suggests that nouns also contribute to the alignment process, most likely by encouraging subjects to engage in the process. It is most likely that the first step in the comparison process is driven by perceptual similarities [20, 25]. However, whether or not later stages of the comparisons, i.e., those that lead to the discovery of more conceptual features (e.g., both are fruits even though they are dissimilar), are driven entirely by these early perceptual associations or rely on higher, independent, conceptually-based inferences remains a topic of considerable debate. It could also be that later, conceptually-based representations of the stimuli (i.e. both belong to the same taxonomic category) are deeply rooted in early perceptual activations or, on the other hand, that there is a conceptual gap between the perceptual activations and deeper conceptual activations. This is a most important issue and one which has not as yet been definitively resolved [54, 55, 56].

Steps and strategies in learning by comparison

The present experiment examines the way comparison situations lead to novel noun generalization in cases where novel words are used with familiar stimuli. Given existing lexical biases, we tackle the question of whether comparisons will lead to conceptually-based decisions, here a taxonomically-related item (e.g. a piece of jewelry in Fig 1.), when a perceptual distractor (e.g. a tyre in Fig 1.) and a thematically-related distractor (e.g. a hand)

are available. Our main question is what visual search strategy, if any, do children follow to reach a conceptually based decision in this comparison design? We will examine this question using eye movement analyses. To date, the available research has focused on the benefits of comparison and on factors that might modulate its effects. To the best of our knowledge, the strategies that children adopt to make sense of the stimuli in a comparison design and generalize a novel object noun have not yet been explored. The aim of our study is to collect eye-tracking data allowing us to reveal the temporal organization of these strategies. Indeed, eye movements provide a precise reflection of the interactions between cognitive processes and the external visual stimuli [57, 58, 59] and can be used to reveal the profile of the search and generalization strategy.

Before we come to our specific predictions, we will first take a look at analogical reasoning tasks which, like novel noun generalization tasks, are also generalization tasks with solution strategies and they have been described in previous studies. The difference is that what is generalized is the relations between items and not the item names. Analogical reasoning tasks are alignment tasks between domains that share the same relational structure [60, 61, 62]. We will therefore examine the past results regarding the strategies used in analogical reasoning tasks and attempt to build further upon these. In analogy tasks, children must consider a setting in which two items – A and B – are related to each other by a specific relation. This setting is often referred to as the “source domain”. Then, in a second, so-called “target domain” that includes an item – C – , children must select an item – D – such that D is to C what B was to A.

Novel noun generalization tasks and analogy tasks are similar since participants must (a) consider the source of information on the initial items (i.e., learning items in novel noun generalization tasks, source domain items in analogical reasoning tasks) in order to extract commonalities; (b) generalize the information they have found to new items (i.e. one item

among multiple generalization options in both the novel noun generalization task and the analogical reasoning task). In our novel noun generalization task, children had to find commonalities between the learning items, extract a representation for the given noun and extend it to the available generalization options (or the target domain items in analogical reasoning tasks). Strategies revealed in the analogical reasoning literature may help us interpret the search patterns observed in our comparison and generalization tasks.

One of the strategies that have been described in analogical reasoning is known as the *Elimination strategy* and involves a one-by-one comparison of the items in the target domain with the items in the source domain, thereby eliminating inappropriate target items until only one is left, which is then selected as the answer [63]. A second strategy is the *Construction strategy*, in which participants start by studying the base pair composing the source domain (A and B) and search for a possible relation that unifies it. They then apply this relation in the target domain until they find a potential D sharing the same relation with C as B has with A [64, 65, 66, 67]. The two strategies can be summarized as follows: Elimination strategy: participants compare learning items and then compare learning items with each available generalization option; Construction strategy: participants compare learning items before applying what they have found to the target domain.

A third strategy has been described by Thibaut and French (2016), the *Task goal-directed strategy* [68]. This strategy refers to the fact that, at the beginning of the trial, younger children look at the set of options from which they have to find the solution (target domain) much more often than older children and adults. Thus, they do NOT systematically compare the A and B items in the source domain as in the construction strategy and nor do they eliminate the options one-by-one by comparing them to A and B. The authors interpreted this pattern as reflecting children's difficulty in inhibiting the main goal of the task, which is "finding the one that goes with C". Whatever the case may be, children and adults have been

found to use different exploration profiles, and this is also the case for both errors and correct answers, thus suggesting how eye-tracking data can help characterize development.

The present study. Search strategies in novel word learning and generalization

The present study seeks to establish the strategies that participants adopt when a comparison design is used in a novel noun generalization task. These strategies will be captured in the form of the temporal organization of gazes up to the point at which children select one stimulus as a referent for the novel noun. To identify the adopted strategies, children have to integrate information coming from various stimuli when the distractors are more salient or as salient as the taxonomic choice. It is therefore important to understand how children organize their searches to reach their final taxonomic decision, that is the temporal dynamics of the novel noun generalization strategies.

The five main aims were to (1) provide a description of novel noun generalization strategies in our comparison design, in terms of patterns of gazes and switches during the trial; (2) interpret this data in terms of a set of strategies; (3) assess the organization of the search and how it interacts with age and task difficulty (i.e., conceptual distance between items); (4) assess whether early gazes predict generalization performance; (5) provide a description of the strategies adopted by good and poor generalization participants.

We tested 5-, 6-, and 8-year-old children in the novel noun generalization task. Six- and eight-year-olds were our two main target groups as we knew [37] that the younger group might still have difficulties with the task, since they were young enough to still be attracted to salient distractors available among the generalization options alongside the taxonomic option. At the same time, we also knew that the older group would primarily give taxonomic answers. Thus, eight-year-olds were expected to produce a majority of mature strategies, whereas six-

year-olds were expected to produce strategies that would reflect a greater influence of the distractors. We also added a five-year-old group after pretests as we became aware that this age group might yield reliable eye-tracking data (see below) and would also testify to young children's difficulties in conceptualizing taxonomic relations, given that it is far more difficult to collect eye-tracking data at such a young age and that this data contains a lot of noise (see [68] for the case of analogies).

We also manipulated the level of difficulty in terms of the conceptual distance between either the learning items themselves or between the learning and generalization options, as conceptual distance might influence the strategy children follow (see below). Indeed, Thibaut and Witt (2017) showed that all conceptual distance conditions differ in difficulty given that it should be easier to find unifying features for conceptually close objects than for distant objects. Here, conceptual distance refers to the taxonomic distance in terms of categorization levels [69]. Following previous studies [39], we define distance in terms of steps in the taxonomic hierarchy: two apples belong to the same basic category and are at distance 0. An apple and an orange belong to the same immediate superordinate category and are at distance 1. Finally, an apple and a piece of meat are at distance 2 as their common level of categorization (food) is separated from them by an intermediate superordinate level of categorization (fruits and meat). In previous work [39], Thibaut and Witt manipulated the conceptual distance between the learning items themselves (e.g., two bracelets versus a bracelet and a watch) and between the learning items and the generalization items (or options) (e.g., a piece of jewelry, near condition, versus a bow tie, distant condition). Results revealed fewer taxonomic choices in the distant generalization condition than in the near generalization condition. Taxonomic generalization at different levels of categorization has also been studied using a single-stimulus design and it has been found that the basic level is the spontaneous naming level rather than the superordinate or subordinate levels of categorization [14, 33, 70].

Overall, our aim is to understand how potential search strategies interact with age and conceptual distance.

In order to reveal these search strategies, we will focus on switches between areas of interest (AOIs) and fixation times on each AOI. Switches reveal comparisons between items and thus the alignment between their features [71]. Fixations reveal which items have been gazed at the most and are thus considered the best sources of information. Hence, the main aims of the analyses are, first, to provide a description of the temporal dynamic of the novel noun generalization strategies. We predict that it will be possible to interpret these strategies in terms of those described above.

An Elimination strategy would involve first comparing the learning items and then going on to compare each of the available generalization options with the learning items, thereby progressively eliminating the generalization items until only one item is left. This should be seen initially in the form of switches between learning items (i.e., comparison of learning items), followed by switches between each of the generalization options and at least one learning item (i.e., search by elimination). The elimination strategy may also be compatible with an even distribution of gazes between the different stimuli, i.e., both the learning items and the options, throughout the entire trial.

A Construction strategy would involve comparing the learning items (characterized by significant switches between learning items). This strategy is most often interpreted in terms of alignment and the identification of common features. The participants should then compare each generalization item with each of the others until one item is chosen as the generalization solution. This strategy will first involve longer periods spent looking at the learning items before the participants then start gazing towards the available generalization options. The Construction strategy predicts very few early gazes towards the options, whereas the Elimination strategy is compatible with early gazes towards the solution set.

These strategies differ mainly in the way the solution set is explored. The Elimination strategy predicts a high proportion of gazes from learning stimuli towards options and very few comparisons between them, whereas the reverse is true for the Construction strategy. The third strategy, namely the Task Goal-directed strategy [68], predicts more switches to and/or longer gazes at the options early in the trial. One way to reveal it is to examine the ratio of switches between the two learning items to switches between either of the learning items and any of the options. Low values on this index at the beginning of the trial should reveal early explorations of the options, that is around the stimuli that are related to the main goal of the task, which is to select one option [44]. We predict an age-related increase in the index value.

Fourth and finally, the associationist views on the importance of shape could lead to an additional strategy in which children, and in particular younger children, first explore the perceptually similar option. This view (Salient Shape view) predicts that the length of the first gazes towards the perceptual distractor should be greater than that of early gazes towards other options or that the number of early switches between learning items and the perceptual distractor should be greater than that of early switches between learning items and other options.

With regard to the role of conceptual distance, we hypothesize that this distance might modulate not only children's generalization performance [44, 49] but also their *strategies*. For example, Bethell-Fox et al. (1984) demonstrated that children shift from one strategy to another as a function of task difficulty, relying on the elimination strategy when difficulty increases, suggesting that the elimination strategy is less costly. Following this logic, we predict that younger children will adopt the elimination strategy more often than older children in the most difficult conceptual distance condition. The reason for this is that the constructive matching strategy is cognitively more demanding since it involves storing the

source item's representation in memory and using it to choose among multiple options without referring back to the source items.

Finally, we will analyze the distribution of early gazes and its relation with the taxonomic decision at the end of the trial and assess whether different patterns can predict which answer a child will choose. Indeed, as shown in an analogical reasoning task [68], the pattern at the beginning of a trial has predictive power. These authors showed that the distribution of gazes in the first third (Slice 1) of a trial predicted whether the trial's outcome would be a correct generalization or an error. Following this logic, we predict that the number of early gazes towards the learning items will increase with age and that a larger number of these gazes will be correlated with a larger number of taxonomic generalizations. In this perspective, we will analyze whether early gaze profiles distinguish between high- and low-achieving participants. We hypothesized that taxonomic choices require a careful analysis of the learning items and fewer early gazes towards the options, meaning that high achievers would have a higher ratio than low achievers of early gazes towards the learning items compared to gazes towards options.

Methods

Participants

Two hundred and thirty-six children subdivided into three age groups were tested individually in a quiet room at their school. Seventy 5-year-old children were recruited (mean age = 63 months; range: 58 – 69), together with ninety-six 6-year-old children (mean age = 80 months; range: 65 – 92), and seventy 8-year-old children (mean age = 100 months; range: 91 – 116). Informed consent was obtained from their school and their parents. These sample sizes were chosen on the basis of Thibaut and French (2016), and Thibaut and Witt (2023), who used smaller samples. We anticipated losing some eye-tracking data in the two younger groups.

The procedure complied with the declaration of Helsinki (1964) and was ethically reviewed and approved within the framework of an official agreement (convention no.: 2019-0679 and endorsement no.: 2020-0566) between the French National Education Ministry's Academic Inspectorate ("Inspection Académique de Côte d'Or"), the University of Bourgogne, and our laboratory.

Materials

We used fourteen familiar categories which were adapted from Thibaut and Witt (2017) (see S1 Table for a full list of materials). An experimental set of seven pictures was built for each category (see Fig 2). This comprised a standard learning item belonging to the targeted category (Fig 2, L₁, e.g. a bracelet), and two other learning items, one conceptually close to the standard learning item (from the same basic level category; Fig 2, L_{2c}, e.g. a curb chain) and one conceptually far from it (from the same superordinate category, Fig 2, L_{2f}, e.g. a watch). Learning items were perceptually similar.

The experimental set also included a taxonomically near generalization item (Fig 2, Ta Near; e.g. a jewel pendant) from the same immediate superordinate-level category as the standard learning item and a distant generalization item (Fig 2, Ta Distant; e.g. a bow tie) from a more remote superordinate category level than that of the standard learning item. Either the near or the distant option was shown during an experimental trial.

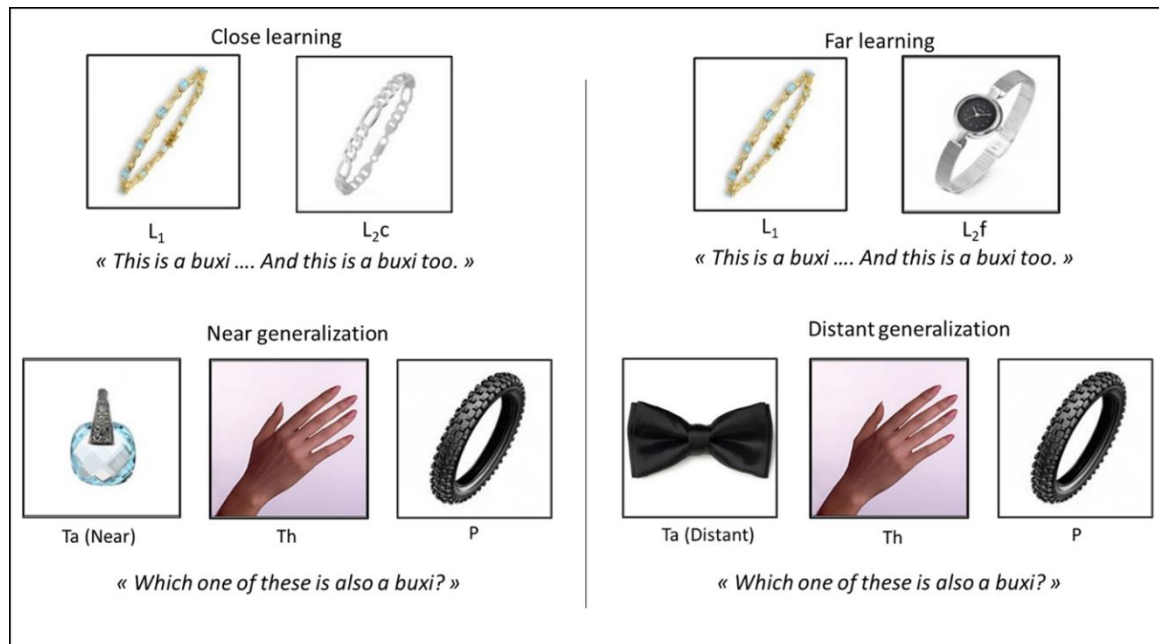


Fig 2. Example of a stimulus set and instructions adapted for experimental conditions resulting from crossing Learning distance (Close vs. Far learning) and Generalization distance (Near vs. Distant).

Finally, two distractors were also included in the set. One was perceptually similar to the learning items but belonged to a semantically unrelated category (Fig 2, P, e.g., a tire). The other was thematically related to the learning objects but not taxonomically or perceptually related to them (Fig 2, Th, e.g., a hand) (See Supporting information 1.A for a full list of materials).

Items were pictures rather than 3D objects because 2D pictures have most often been used in novel noun generalization studies and because they are better suited for studying children's strategies on the basis of eye-tracking measures. It might be thought that 2D pictures are not as informative as 3D items in terms of perceptual input. One question was whether the perceptual distractor would attract sufficient attention, especially in the case of familiar stimuli. However, we built on the experiment of Thibaut and Witt (2023), who used

the same stimuli in a no-comparison condition. As expected, children selected the perceptual match above chance level, as predicted by a shape bias.

Independent similarity ratings from 54 university students (mainly psychology students, who participated in return for course credits, mean age = 20.15 years; 47 females) were recorded to control the material. We first controlled for the differences in the conceptual distance between conceptually related items. Conceptual similarity ratings confirmed that the close learning stimuli were conceptually closer to their related standard learning stimulus than the far learning stimuli, (scale 1-7, $M_{Close} = 5.52$, 95% CI [5.32, 5.72] ; $M_{Far} = 4.44$, 95% CI [3.89, 4.99]), $t(13) = 3.31$, $p < .006$, $d = 0.89$ (Bonferroni-corrected p-value threshold $.05/8 = .00625$), and that near generalization stimuli were conceptually nearer the learning stimuli than distant generalization stimuli, (scale 1-7, $M_{Near} = 3.75$, 95% CI [3.45, 4.05]; $M_{Distant} = 2.48$, 95% CI [2.06, 2.90]), $t(13) = 4.41$, $p < .006$, $d = 1.81$.

The taxonomically related generalization options have to be perceptually less similar to the learning items than the perceptual distractors. Perceptual similarity ratings revealed that the perceptual choices were perceptually more similar to the learning items than the taxonomic choices in both the close and the far learning conditions (scale 1-7, $M_{Perceptual} = 4.77$, 95% CI [4.49, 5.05]; $M_{Near} = 2.13$, 95% CI [1.81, 2.45]; $M_{Distant} = 1.86$, 95% CI [1.550, 2.21]), $t(13) = 10.59$, $p < .001$, $d = 2.83$, and $t(13) = 11.63$, $p < .001$, $d = 3.11$.

We also tested whether the thematically related items were semantically related to the learning items in both close and far learning conditions (scale 1-10, $M_{semantic} = 7.74$, 95% CI [7.34, 8.14]; $M_{Near} = 6.30$, 95% CI [5.38, 7.22]; $M_{Distant} = 4.08$, 95% CI [3.07, 5.09]), $t(13) = 2.95$, $p < .05$, $d = 0.79$, and $t(13) = 6.62$, $p < .001$, $d = 1.77$.

Importantly, we also performed perceptual similarity and conceptual similarity ratings between the close learning stimuli (e.g., two apples) and the far learning stimuli (e.g., an apple and a cherry), on the one hand, and the taxonomically related generalization item, on the

other. They showed that, overall, the generalization stimuli were equally distant, both perceptually and semantically, from both types of learning item. This was true for both types of generalization item. Near generalization items: perceptual distance (scale 1-7, $M_{Close} = 2.18$, 95% CI [1.82, 2.54]; $M_{Far} = 2.02$, 95% CI [1.71, 2.33]), $t(13) = 1.58$, $p = .14$, $d = 0.42$, conceptual distance (scale 1-7, $M_{Close} = 3.09$, 95% CI [2.82, 3.36]; $M_{Far} = 3.36$, 95% CI [3.05, 3.67]), $t(13) = -2.96$, $p = .01$, $d = -0.80$ (Bonferroni-corrected p-value threshold $.05/8 = .006$). Distant generalization: perceptual distance (scale 1-7, $M_{Close} = 1.88$, 95% CI [1.56, 2.20]; $M_{Far} = 1.86$, 95% CI [1.47, 2.25]), $t(13) = .24$, $p = .81$, $d = 0.06$, conceptual distance (scale 1-7, $M_{Close} = 2.18$, 95% CI [1.82, 2.54]; $M_{Far} = 2.02$, 95% CI [1.71, 2.34]), $t(13) = -.19$, $p = .85$, $d = -0.05$. This was crucial because we wanted to avoid the possibility that performance differences between near and distant generalization items might be due to perceptual or semantic similarity differences between them and the learning items. For example, if we observed a difference between near and distant generalization items (e.g. between jewelry pendant and bow tie), we did not want it to be due to semantic information (e.g., the fact that the pendant is more thematically related to the bracelet than the bow tie) other than the difference in conceptual distance (See S2 Table for details of the ratings).

Fourteen different bisyllabic labels (pseudo-words) [72], which are easier to remember than monosyllabic pseudo-words, were used to name the learning objects in each trial. Syllables were of the CV type, which is the dominant word structure in French (from Lexique.org, bari, buxi, daxo, jito, malan, missi, muno, nati, peco, pina, rula, sefu, soki, vira, youma, zatu [73])

Two learning items were used in each trial, either the standard learning item and the close learning item ($L_1 - L_{2c}$, close learning condition) or the standard learning item and the far learning item ($L_1 - L_{2f}$, far learning condition). One of the two generalization items was used (near or distant item according to the generalization condition) together with the

perceptually and the thematically related distractors. The two learning items were displayed side-by-side at the top of the screen and the three generalization options were displayed side-by-side below the learning items (see Fig 2). The learning items' positions were randomized, as were the positions of the generalization items. For each trial, the pictures were displayed simultaneously on a Tobii T120 eye-tracker device with a 1024x768 screen resolution until the answer was given.

Each picture constituted an Area of Interest (AOI) for the eye-tracking analysis. Picture size was 500x500pixels and each picture was surrounded by a square black outline. The boundaries of the AOIs corresponded to a picture's outline. A standard fixation cross was shown for 3 seconds between each trial. Each experimental session started with a standard calibration phase conducted after the three warm-up trials. The experiment was run with E-prime® software (Psychology Software Tools, Pittsburgh, PA).

Children's general knowledge and language development were tested using the EVIP (Echelle du Vocabulaire en Images de Peabody) standardized French vocabulary test. As expected, children's scores improved with age, but no significant correlations were found between vocabulary scores and generalization scores (Pearson partial correlation $r = 0.07$, $p = 0.45$)

Procedure

The child and the experimenter were seated at a table in a quiet room in the child's school. The experimenter introduced the experiment as a game to be played with a bear named Sammy. "Hello, we are going to play together, and we are going to play with a bear called Sammy. Look, this is Sammy, he lives far away from here and speaks a different language; he does not speak like us and we are going to learn his language."

It was critical to limit external elements that might have influenced the children's gaze patterns during the task (e.g., hand movements by the experimenter, sequential item appearance on the screen etc.). To eliminate the effect of the item's appearance, all of the trial's stimuli always appeared simultaneously on the screen, thus making stimulus location (learning and option items) predictable. When the items appeared, the experimenter gave the trial instruction: "*See, Sammy's mummy says this is a buxi. And this is a buxi too. Sammy must find another buxi. Can you show which one of these three is also a buxi, to help Sammy? Can you point to the other buxi?*" To eliminate any effect of the experimenter's movements, he did not make any hand movements towards the screen during testing. To ensure that this did not affect children's understanding of the task, a warm-up phase, intended to explain the task to the children, was conducted before eye-tracking recording started.

In the warm-up phase, children saw three trials which were identical to those seen later during the test phase except for the instruction given. In the first warm-up trial, the experimenter added location information and pointing movements to the items he was talking about. "*See, Sammy's mummy says this item at the top of the screen is a buxi. And this one next to it is a buxi too (Exp points to them). Sammy must find another buxi. Can you show which one is also a buxi, to help Sammy? Can you point to the other buxi among these items at the bottom of the screen (pointing at the three options on the screen)?*" In the second warm-up trial, the experimenter gave the same instruction without pointing and without location information (e.g., "at the top of the screen"). In the third warm-up trial, the experimenter gave the same instruction as in the test trials, still without pointing. If the child did not answer in the second warm-up trial, the experimenter repeated the instruction together with the pointing movements. It was decided that children who did not answer in the third warm-up trial would be excluded because we could not be certain that they had understood the task. However, no

child was excluded at this stage. After the warm-up trials, the experimenter started the eye-tracking calibration phase, which was followed by the experimental trials themselves.

During the test phase, the experimenter showed the child 14 trials as described above. The experimenter used no hand movements or other ways of highlighting the relevant items. The child chose an item by pointing to it and the experimenter selected it with the mouse. Response times were recorded from the time when the item appeared on screen to the experimenter's click, which also ended the trial and cleared the display.

The order of the fourteen experimental trials was randomized, as were the names used. A blank screen with a central fixation cross was shown for 3 seconds between each trial. Participants knew the items taken from Thibaut & Witt (2017; 2023), who tested 4- and 5-year-old children with these items.

Eye-tracking data were recorded as follows. The Tobii 120 eye-tracking device, equipped with two cameras, recorded the position of the two eyes every 8.33ms. Raw data consisted of all spatial coordinates of eye positions throughout trials [74].

Results

Performance data

Performance was measured as the percentage of taxonomic generalizations (i.e., percentage of taxonomic choices) and children's reaction times in milliseconds. The main aim of the present study was to understand the strategies that lead to taxonomic generalization. Following standard practice, we initially focused on correct (i.e., taxonomic answers) trials only, because we wanted to describe search strategies that lead to a correct answer. It is more difficult to interpret incorrect answers as they may result from a real choice or from random

choices when children are inattentive (see below for a comparison of high-achieving and low-achieving participants).

Generalization scores

First, we analyzed the percentages of taxonomic choices. We ran a three-way repeated-measures ANOVA on participants' percentage of correct answers, with age (5, 6, 8 years old) and learning distance (close, far) as between-factors and generalization distance (near, distant) as a within-factor. The analysis revealed an effect of generalization distance: $F(1, 179) = 54.74, p < .001, \eta_p^2 = .23$, age: $F(2, 179) = 9.95, p < .001, \eta_p^2 = 0.1$ and an interaction between generalization distance and age: $F(2, 179) = 3.30, p < .05, \eta_p^2 = .04$ (see Fig 3). All other main effects and interactions failed to reach significance ($p < .05$).

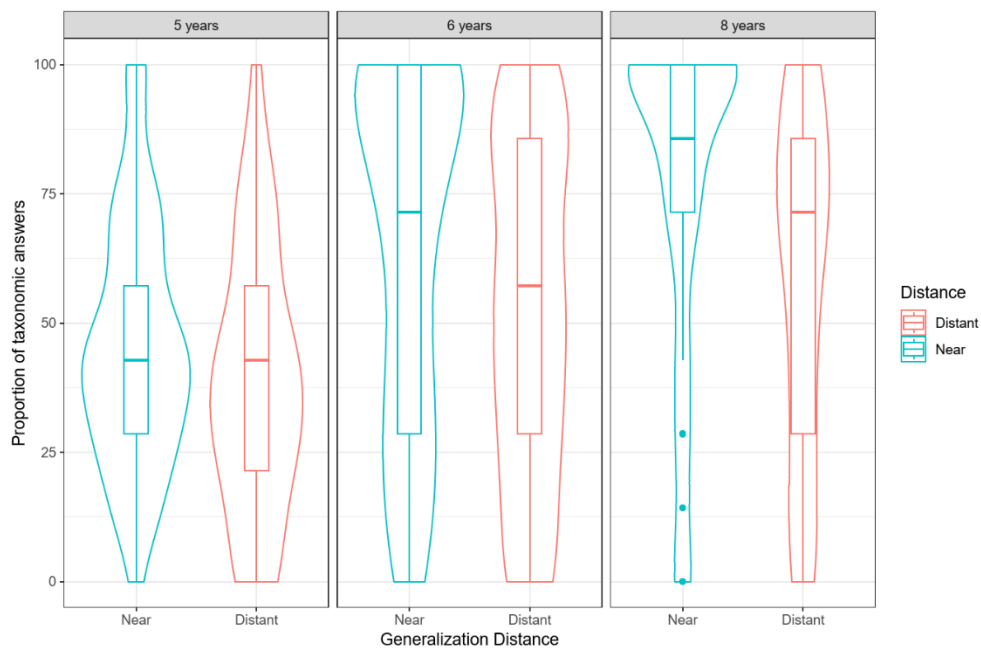


Fig 3. Percentage of taxonomic choices as a function of age (5, 6 and 8 years old) and generalization distance (near, distant).

For 6- and 8-year-old children, an a posteriori Tukey analysis revealed that near generalization scores were significantly higher than distant generalization scores ($p_s < .001$).

There was no significant difference between near and distant scores for 5-year-old children. One-sample T-tests revealed that all groups scored above chance (set at 33%), except for 5-year-old children, who scored at chance in distant generalization. Near generalization scores tested to chance (.33): $t(184) = 3.58, p < .001$. Distant generalization scores tested to chance (.33): $t = 1.43, p = .16$. 6 years. Near generalization: $t(184) = 8.50, p < .001$; Distant generalization: $t(184) = 5.34, p < .001$. 8 years. Near generalization: $t(184) = 10.30, p < .001$; Distant generalization: $t(184) = 5.71, p < .001$.

Reaction times

We ran a three-way repeated-measures ANOVA on children's average reaction times for trials that led to correct answers, with age (5, 6, 8 years old) and learning distance (close, far) as between-factors and generalization (near, distant) as a within-factor. The analysis revealed an effect of age, $F(2, 182) = 7.58, p < .001, \eta_p^2 = .08$, and interactions between learning distance and age, $F(2, 179) = 5.38, p < .05, \eta_p^2 = .06$ (see Fig 4), and between generalization distance and age, $F(2, 182) = 3.42, p < .05, \eta_p^2 = .04$ (see Fig 5). Five-year-olds were slightly faster in the distant condition, but again performed at chance and might therefore have failed to perform the task in this condition. An a posteriori Tukey analysis revealed that 6-year-old children's reaction times were significantly faster in close than in far learning trials ($M_{close} = 10,208$ ms, $M_{far} = 12,307$ ms; $p < .05$). An a posteriori Tukey analysis revealed that reaction times decreased with age between 6 and 8 years (5- vs 6-year old children: $p_{Near} = .66, p_{Distant} = .95$; 6- vs 8-year old children : $p_{Near} < .05, p_{Distant} < .05$), showing that children find the task easier with age.

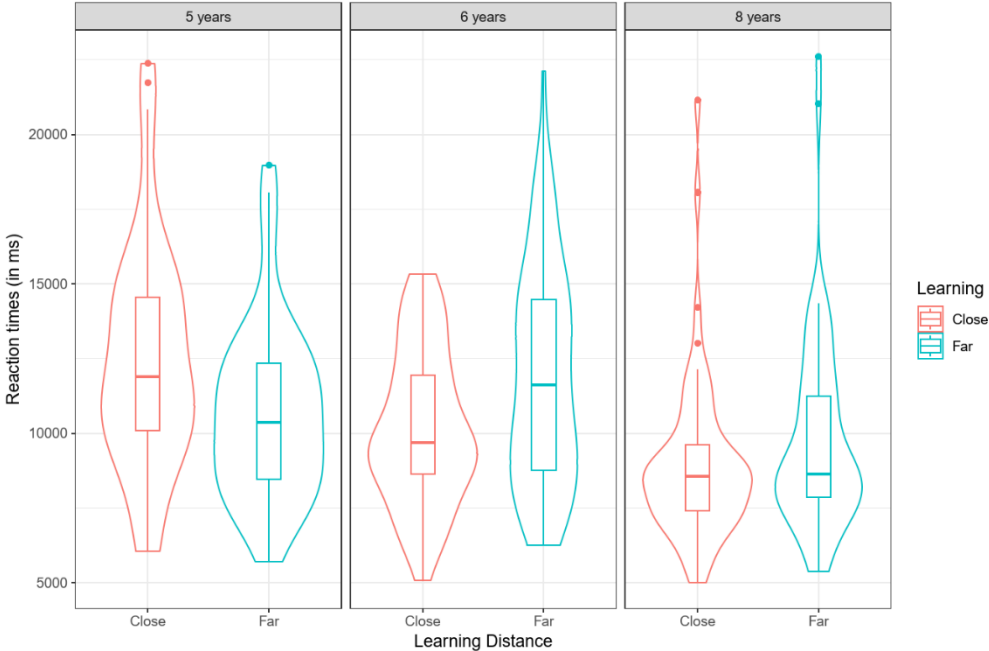


Fig 4. Reaction times (ms) for correct answers as a function of learning distance (close, far) and age (5-, 6-, 8- years-old).

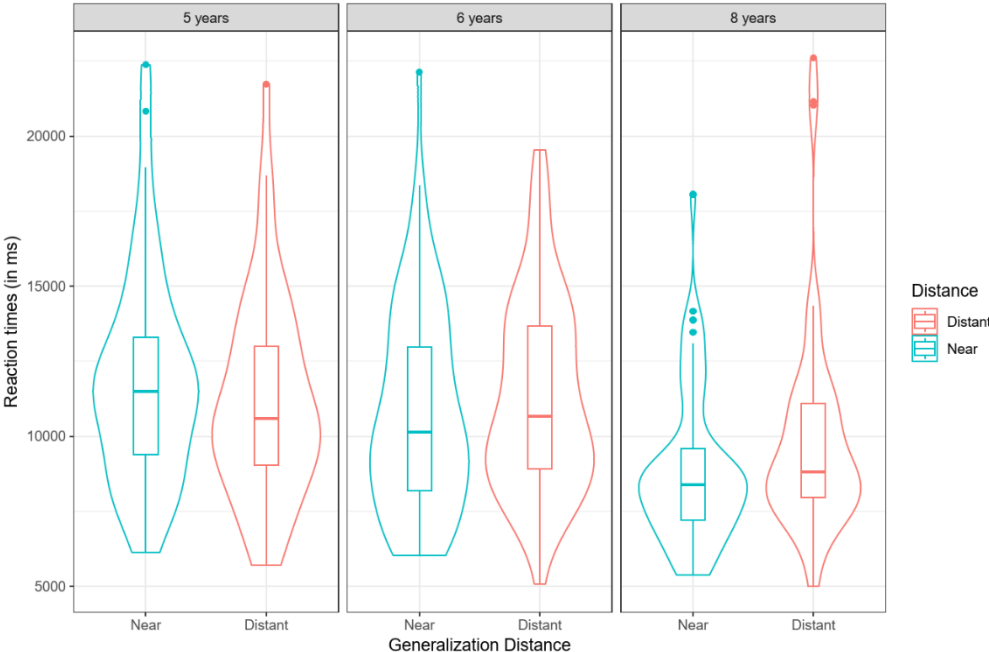


Fig 5. Reaction times (in ms) for correct answers as a function of age (5, 6, 8-year-old) and generalization distance (near, distant).

Taken together, these results show a consistent pattern of results in which the distant generalization condition was the most difficult condition, in terms of both RTs and the percentage of taxonomic choices, thus confirming recent performance data reported by Thibaut and Witt (2023).

Eye-tracking data: strategy description

Learning distance

Learning distance did not have an effect on the percentage of taxonomic generalization choices and its size effect on reaction times was small. For these reasons, and after verifying that learning distance did not interact with other factors in the eye-tracking analyses, we excluded this factor from further analysis to simplify the design and improve the robustness of our results. Generalization distance was not excluded from the analyses.

We verified the absence of an effect of learning distance in a five-way repeated-measures ANOVA on the proportion of fixation times or the log-transformed number of switches (see below), with learning distance (close, far) and age (5, 6, 8) as between-subject factors and generalization distance (near, distant), time slice (beginning, middle, end) and AOI (L, Th, Ta, P) or switch type (LL, LTh, LTa, LP, ThTaP) as within-subject factors.

In the case of fixation times, learning distance appeared as a factor in the interactions between learning distance, time slice and AOI, $F(6, 714) = 2.32, p < .05, \eta_p^2 = .02$, and learning distance, time slice, generalization distance and AOI, $F(6, 714) = 2.12, p < .05, \eta_p^2 = .01$. These interaction's effect sizes were small and post hoc tests did not reveal any differences between close and far learning groups (see also Thibaut & French, 2016, who also removed non-interacting factors and factors involved in small-sized interaction effects, for a discussion.) With regard to the analyses of switches, learning distance did not appear as a factor in any of the interactions.

Collection and processing of eye-tracking data

Data from seven children were removed because eye-tracking calibration was unsuccessful (1 five-, 4 six- and 2 eight-year-old children) due to the fact that inadequate contrast between eye color, room lighting and screen lighting made the pupil too difficult to detect. Data from a further twelve children was incomplete because they interrupted the task or because eye-tracking calibration was interrupted during the task because the children were unable to stay still (ten 5- and two 6-year-old children). This data set was also excluded from further analyses.

We also excluded trials for which more than 50% of eye positions had been lost and we excluded subjects for whom more than 50% of trials had been excluded. This resulted in the exclusion of 32 subjects (twelve 5-year-olds, ten 6-year-olds; and ten 8-year-olds). Among the remaining participants, we lost an average of 1.7 trials for 5-year-old children, 1.1 trials for 6-year-old children and 1.1 trials for 8-year-old children.

Eye-tracking data was divided into three equal time slices. The length of a time slice was equal to a third of a trial's total fixation time. We first focus on gazes towards AOIs and switches (transitions). Gaze duration (or looking times) for areas of interest (i.e., AOIs) tells us which stimuli were attended to and for how long while solving the problem and also reveal the depth of processing of the item. Switches between items (or saccades or transitions) tell us which items were compared and can be interpreted as an attempt to find commonalities between X and Y [68]. The two measures reveal different aspects of the search, as a participant could study an item for a large amount of time without comparing it with other items, while gaze duration, unlike transitions, tells us nothing about which items were compared during trials.

To meet the normality requirements for ANOVAs, we applied normal transformations to our data. Log transformations were applied to the number of switches. We applied Greenhouse-Geisser or Huynh-Feldt corrections to F scores when sphericity scores required them.

Fixation times

First, we analyzed the proportion of fixation times allocated to the AOIs. We analyzed the proportion of fixation times during trials. Following standard practice [74, 75], a gaze towards an AOI was considered as a fixation when both eye positions were recorded inside the AOI's boundaries for longer than 60ms. Because of sample size was large we checked that the data was normally distributed with Q-Q plots by group that confirmed that proportions of fixation times met the normal distribution requirements for ANOVAs.

We ran a four-way ANOVA on the proportion of fixation times in correct trials, with age (5, 6, 8 years) as a between-factor and generalization distance (near, distant), time slice (beginning, middle, end) and AOI (L, Th, Ta, P) as within-factors. The analysis revealed a simple effect of AOI, $F(3, 366) = 74.50, p < .001, \eta_p^2 = .38$. The analysis also revealed interaction effects between generalization distance and time slice, $F(2, 244) = 3.27, p < .05, \eta_p^2 = .03$, and between time slice and AOI, $F(6, 732) = 85.54, p < .001, \eta_p^2 = .41$ (see Fig 6.). A three-way interaction was also revealed between age, time slice and AOI, $F(12, 732) = 2.40, p < .05, \eta_p^2 = .04$.

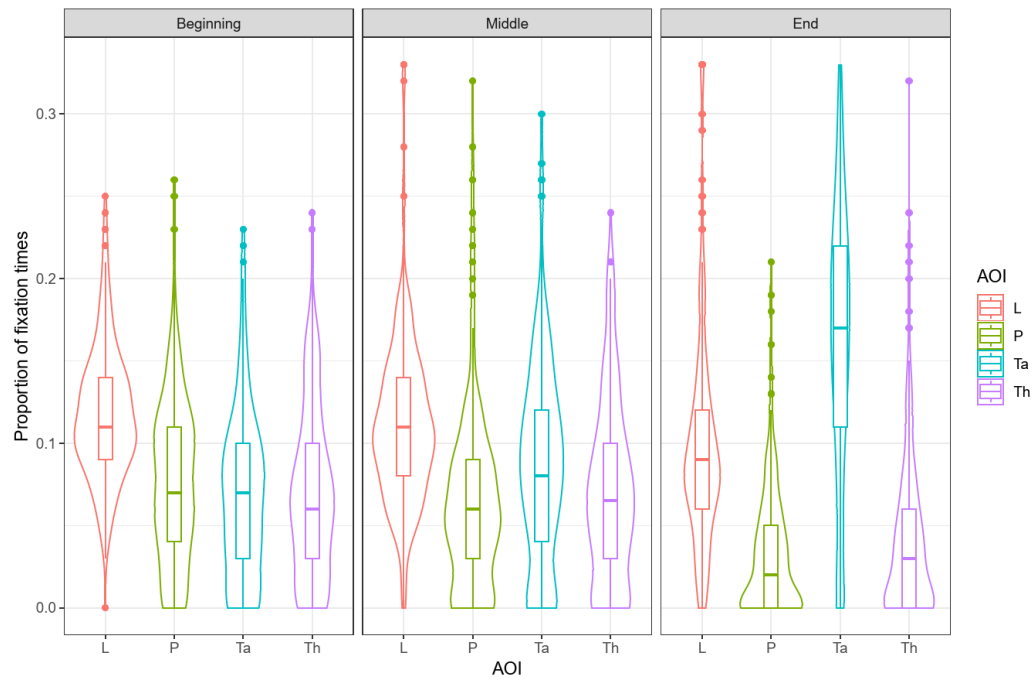


Fig 6. Proportion of fixation times for correct answers as a function of time slice (beginning, middle, end), and Area of Interest (L, Th, Ta, P).

The triple interaction had a small effect size ($\eta^2_p = .04$). A post hoc Tukey test revealed no significant difference between age groups. All the observed differences were between AOI and time slice and were redundant with those revealed by the two-way interaction. To simplify, we therefore focused on the two-way interaction between time slice and AOI (see Fig 6). We ran contrast analyzes using Bonferroni corrections to analyze this interaction. The contrasts of interest were those between different AOIs during one and the same time slice and those between one and the same AOI across different time slices. The Bonferroni-adjusted α -value threshold was set at $p = .0016$.

The interaction reveals that at the beginning of a trial, children looked significantly more at the learning items than at the other items ($L > Th$ $p < .001$, $L > Ta$ $p < .001$, $L > P$ $p < .001$). In the middle slice, they still looked more at the learning items ($L > Th$ $p < .001$, $L > Ta$ $p < .001$, $L > P$ $p < .001$). At the end of a trial, they looked more at the taxonomic option ($Ta > L$ $p < .001$, $Ta > Th$ $p < .001$, $Ta > P$ $p < .001$) and also looked more at the learning

items than the distractors (End: $L > Th$ $p < .001$, $L > P$ $p < .001$). Taken together, these contrasts show that children started by looking at the learning items and continued to do so throughout the whole trial. While they also looked at all the available generalization options at the beginning of the trials, they gradually turned away from the distractors (i.e., Th and P) in order to concentrate their attention on the taxonomic option.

Contrasts for each AOI across slices reveal that children's looking times to the learning items decreased significantly between the middle and the end of the trials (L: Middle > End $p < .001$). The time spent looking at the thematically related distractor decreased between the middle and the end of each trial; it increased for the taxonomic item and decreased for the perceptual distractor as the trial progressed (Th: Middle > End $p < .001$; Ta: Beginning < Middle $p < .001$, Middle < End $p < .001$, P: Beginning < End $p < .001$). These contrasts confirm that, as each trial progressed, children's attention moved away from the distractors and towards the taxonomic item.

Switches

Following common practice [74, 75], switches were defined as movements from one AOI to another, provided that children's gaze stayed focused on each AOI for at least 60 ms. Since there were 5 stimuli, which could a priori result in 20 different switches, we aggregated them in five different switch types that were the most relevant for our analysis. Thus, LL switches are all switches between L_1 and L_2 , LTh switches are switches between L_1 or L_2 and Th and vice versa. Similarly, LTa switches are switches between L_1 or L_2 and Ta and vice versa; LP switches are those between L_1 or L_2 and P and vice versa; and ThTaP switches are all switches between the available generalization options (Th, Ta and P).

We analyzed the log transformation of the number of switches to fit the normality assumptions for ANOVAs. We ran a four-way ANOVA on data from correct trials, with age

(5, 6, 8 years) as a between-factor and generalization distance (near, distant), time slice (beginning, middle, end) and switch type (LL, LTh, LTa, LP, ThTaP) as within-factors. The analysis revealed simple effects of age, $F(2,108) = 8.82, p < .001, \eta_p^2 = .14$, generalization distance, $F(1,108) = 24.03, p < .001, \eta_p^2 = .18$, time slice, $F(2,213) = 73.09, p < .001, \eta_p^2 = .40$, switch type, $F(4,432) = 40.92, p < .001, \eta_p^2 = .28$. The effects of age and generalization distance showed that 5-year-old children made less switches than 6- and 8-year-old children ($M_5 = 0.34; M_6 = 0.46; M_8 = 0.47; 5 < 6, p < .001$) and that children made less switches in the distant generalization setting than in the near generalization setting ($M_{\text{Near}} = 0.46; M_{\text{Distant}} = 0.38; p < .01$).

The most interesting result involving switch type and time slice was the interaction effect between time slice and switch type, $F(7,789) = 11.24, p < .001, \eta_p^2 = .09$, as it reveals the temporal dynamics of children's searches across switches (see Fig 7).

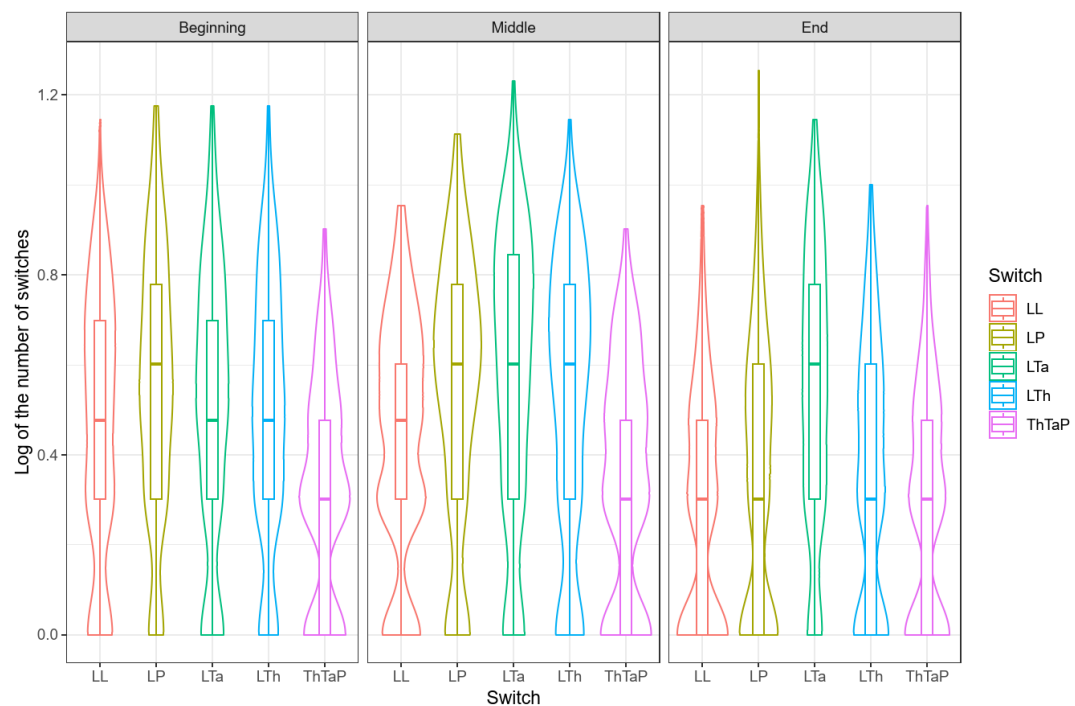


Fig 7. Log of the number of switches for correct trials as a function of time slice (beginning, middle, end), and switch type (LL, LTh, Lta, LP, ThTaP).

We ran contrast analyses with a Bonferroni correction set at $p < .0011$. The contrasts of interest were those between different switch types during one and the same time slice, and those between the same switch types over different time slices. The interaction reveals that, at the beginning of a trial, there were significantly more LP switches than LL switches ($p < .001$) and more LTh switches than LL switches ($p < .001$), showing that children's attention was initially attracted to the distractors.

At the beginning of each trial, children also made less ThTaP switches than all the other switch types ($p < .001$). In the middle, there were more switches between learning items and available generalization options (i.e., LTh, LTa, LP) than between learning items themselves (i.e., LL) (LTh > LL $p < .008$, LTa > LL $p < .001$, LP > LL $p < .001$). There were also fewer switches between available generalization options (i.e. ThTaP) than any other type of switch in the middle of the trials (ThTaP < LL-LTa-LTh-LP, $p < .001$). At the end of the trials, children made significantly more LTa switches than all the other switch types, which did not differ from one another (LTa > LL $p < .001$; LTa > LTh $p < .001$; LTa > LP $p < .001$; LTa > ThTaP $p < .001$).

These contrasts clearly show that children started by comparing all the items with one another, showing a slight preference for LP switches at the beginning of the trials, likely due to the salience of the perceptual distractors. While they then continued to make comparisons between learning items (i.e., LL switches) throughout each trial, they made less switches towards the distractors (i.e., LTh and LP) in the middle and at the end of the trials. This shows that they turned away from these distractors and gradually made more LTa switches.

Individual contrasts run for each switch type across time slices reveal that the number of LL switches decreased at the end of each trial (Middle > End, $p < .001$). Not surprisingly, the same switching time-course was found for LTh switches (Middle > End $p < .001$) and for LP switches (Middle > End $p < .001$). These contrasts show that children systematically

explored the three options in the middle of each trial but turned away from the distractors during the trial as they reached a taxonomic decision.

First gazes: distribution and predictive power.

The first two analyses reveal that all AOI were gazed at and that all switch types were produced in the first time slice. The fact that all the switch types appeared in the first time slice is probably due to the small number of stimuli and the simplicity of the task, which allowed participants to explore the entire task space during the first few seconds. Hence, the pattern observed during the first slice might have obscured regularities in gaze sequences between learning items and between learning items and options in the early steps of the trial. For this reason, we analyzed the nature of the first five fixated items. Given our research hypotheses, we considered whether children would first gaze at the learning items, aligning them in order to extract their common properties, or whether they would first look at the generalization items in search of the solution. We hypothesized that more gazes towards the learning items rather than the options would be predictive of correct trials. To test this hypothesis, we calculated a ratio measuring the orientation of participants' attention for each of the first five gazes. This ratio was defined as the number of gazes towards the learning items made by a participant divided by the number of gazes directed towards the generalization items by the same participant and was named Learning-items/Generalization-items (hereafter Learn/Gen). By definition, higher values of the ratio mean that children looked more at the learning items, and that a participant focused more on the learning items for this particular gaze rank (1st, 2nd, etc.). This kind of integrative measure is very useful as it aggregates the strategy adopted by participants under a single value and powerful hypotheses can be derived from it. First, if our hypothesis that participants will first fix their gaze on learning items is highly predictive of performance [68, 76], we predict that older children will

have higher ratio values for the initial gazes (e.g., 1st and 2nd) and that higher values for these gazes will be correlated with a higher number of taxonomic choices.

We ran a three-way repeated-measures ANOVA on the Learn/Gen ratio with age (5, 6, 8-years old) as a between-factor and generalization distance (near, distant) and gaze position (1, 2, 3, 4, 5) as within-factors. No corrections to the data were needed. The analysis revealed effects of age, $F(2, 121) = 3.20, p < .05, \eta^2_p = .05$ ($M_5 = 1.01, M_6 = 1.31, M_8 = 1.26, 5yo < 6yo p < .05$), generalization distance, $F(1,121) = 8.02, p < .05, \eta^2_p = .06$ ($M_{Near} = 1.29 ; M_{Distant} = 1.10$), and gaze position, $F(4, 484) = 17.56, p < .001, \eta^2_p = .13$ ($M_1 = 1.76 ; M_2 = 0.94; M_3 = 1.18; M_4 = 0.96; M_5 = 1.14$). Post hoc Tukey tests reveal that the Learn/Gen ratio was significantly higher for the first gaze than for each of the other gazes ($p < .001$). For the following ratios (gaze 2 to 5), there was no significant difference between gazes. No interaction reached significance ($p > .05$). Thus, the ratio is seen to increase with age, a finding which is compatible with our hypothesis that the ability to make comparisons increases with age. We also found that the ratio decreased after the first gaze, meaning that children first looked at the learning items but then was distributed equally between the learning items and the options.

We next tested whether early attention orientation was predictive of children's taxonomic generalization performances [39, 68]. To this end, we first computed Spearman correlations, for correct trials, between participants' Learn/Gen ratios and their percentage of taxonomic answers. The correlations revealed a positive correlation between the ratio value and the number of correct trials for the first five gazes (see S3 Table).

We then used general linear models with children's generalization scores (i.e., proportion of taxonomic generalizations) as outcome and Learn/Gen ratios as predictors, while controlling for age and learning distance. Our best-fit models (M3 for near generalization, M4 for distant generalization) included the Learn/Gen ratio for the first three

gazes in near generalization cases (see Table 1), and the Learn/Gen ratio for the first four gazes in distant generalization cases (see Table 2). We ran ANOVAs on the best fitting models (see S4 Table and S5 Table).

These results strengthen our hypothesis that an initial focus on the learning items, and therefore not studying the options before interpreting the learning items, predicts taxonomic generalization in both near and distant generalization conditions. It is interesting to note that successful generalization in the distant generalization condition was predicted by a longer string of gazes towards learning items (four gazes) than in the near generalization condition (three gazes). This is consistent with the idea that children needed to compare the learning items more frequently in the more difficult condition in order to successfully align the learning items, with the result that longer strings remained predictive.

Table 1: The goodness of fit of the generalized linear models, with children's generalization scores (in near generalization cases) as the outcome and Learn/Gen ratios as predictors.

The model controls for age and learning distance.

Outcomes	Model	Predictors	<i>Df</i>	AIC	R ²	Adjusted R ²	<i>p</i> -value
Percentage of correct answers	M0	Age + Learning	1	1338	.163	.151	
	M1	Age + Learning + Learn/Gen ₁	1	1302	.362	.348	<.001
	M2	Age + Learning + Learn/Gen ₁ + Learn/Gen ₂	1	1295	.401	.383	<.01
	M3	... + Learn/Gen ₃	1	1292	.420	.399	<.05
	M4	... + Learn/Gen ₄	1	1293	.425	.418	.29

Table 2: The goodness of fit of the generalized linear models, with children's generalization scores (in distant generalization cases) as the outcome and Learn/Gen ratios as predictors.

The model controls for age and learning distance.

Outcomes	Model	Predictors	<i>Df</i>	AIC	R ²	Adjusted R ²	<i>p</i> -value
Percentage of correct answers	M0	Age + Learning	1	1207	.058	.043	
	M1	Age + Learning + Learn/Gen ₁	1	1158	.370	.355	<.001
	M2	Age + Learning + Learn/Gen ₁ + Learn/Gen ₂	1	1152	.407	.388	<.01
	M3	... + Learn/Gen ₃	1	1143	.456	.433	<.001
	M4	... + Learn/Gen ₄	1	1139	.484	.458	<.05
	M5	... + Learn/Gen ₅	1	1138	.495	.466	.102

Different strategies for different performance? High achievers versus low achievers.

The previous analysis revealed that initial gazes were highly predictive of children's scores, and in particular that higher Learn/Gen ratio values predicted a higher percentage of taxonomic responses. However, the fact that this ratio for earlier gazes significantly predicted the percentage of correct generalizations does not mean that high ratios necessarily equate to high performance levels. The present analysis complements the previous one. Our general hypothesis here is at the child level and considers that higher-performing children also follow different strategies that lead to the chosen option, either correct (taxonomic) or incorrect (perceptual or thematic). This means that the choice they eventually make depends on what they do throughout the entire trial, before coming to their decision. Indeed, it is also possible that both correct and incorrect choices are prepared for by the early search steps. Early gazes

and/or switches have been shown to reliably predict the outputs (error or correct response) from analogy tasks [68, 76]. On the other hand, errors might result from the fact that children finally opted for one of the two non-taxonomically related choices at the decision stage despite having followed the same search strategy in both correct and incorrect trials.

Following this logic, we tested whether high achievers and low achievers followed the same sequence of gazes at the beginning of the trial in order to confirm (or disconfirm) the importance of early comparisons for correct generalization. Indeed, the last steps in the decision-making process are less interesting since, by definition, low achievers choose more distractors than high achievers and therefore gaze more at them and less at the taxonomic match.

For this reason, we now focus on high and low achievers' search profiles for the first five gazes as in the previous analysis. High-achieving children were defined as those who obtained eleven or more correct answers (over 14 trials) and low-achieving children as those obtaining four or less correct answers. We analyzed 6-year-old children's data only, since 5-year-old children had low scores with too few "high-achieving" children, whereas the reverse was true for 8-year-old children. Among the 6-year-old children, there were 30 high-achieving children and 18 low-achieving children. This measure focuses on two highly contrasted groups in terms of performance. It also ignores participants characterized by more ambiguous profiles, who were also likely to exhibit less-well defined strategies.

We analyzed the Learn/Gen ratio described above. We hypothesized that high-achieving children would focus more on the learning items at the beginning of the trial, that is would have a higher Learn/Gen ratio than low-achieving children. We ran a 3-way repeated-measures ANOVA on the Learn/Gen ratio computed on the first five gazes for correct answers, with subject profile (low, high) as a between-factor, and generalization distance (near, distant) and gaze position (1, 2, 3, 4, 5) as within-factors. Results revealed significant

effects of subject profile, $F(1, 42) = 35.4$, $p < .001$, $\eta^2_p = .46$, with low-achieving subjects' ratios being significantly lower than those of high-achieving subjects, $M_{Low} = 0.80$, $M_{High} = 1.61$. They also revealed a main effect of gaze position, with the Learn/Gen ratio being significantly higher for the first than for the second and fourth gaze positions, $M_1 = 1.79$, $M_2 = 0.92$, $M_3 = 1.07$, $M_4 = 0.92$, $M_5 = 1.32$, (t-test with Bonferroni-corrected significance level set at $p = .005$, $p < .005$). The reason why M_2 and M_4 were lower than M_1 but not M_3 and M_5 is not entirely clear. Starting with a learning item, some participants gazed at another learning item, whereas others tended to gaze back and forth between learning and transfer items. No other effect reached significance ($p > .05$).

These two effects reveal that high-achieving children focused significantly more than low-achieving children on learning items than on the possible options during their initial gazes, confirming previous results [68]. This result also makes sense in terms of the strategies: children who followed our third strategy, the goal-directed strategy described by Thibaut and French [68], would have subsequently been at risk of choosing the perceptual distractor, rather than the option resulting from a careful alignment of the learning items, namely the taxonomic option.

Discussion

Building on earlier results on novel noun learning obtained by Gentner and Namy, who used a comparison design, we proposed that this ability can be more generally described in terms of the alignment of the learning stimuli. Our main goal was to use an approach similar to that adopted in other analogical reasoning studies [68, 76, 77, 77] in order to provide a description of the temporal dynamics of children's learning/generalization strategies when making comparisons. The main aims were to (1) provide a description of novel noun generalization strategies in a learning-with-comparison design in terms of patterns of gazes

and switches during the trial; (2) interpret this data in terms of a set of hypothetical strategies; (3) assess the organization of the search and how it interacts with age and task difficulty (i.e., conceptual distance between items); (4) assess whether early gazes predict generalization performance; (5) provide a description of strategies that lead to high performance (good achievers) or poor performance (low achievers). Overall, our results reveal that the three age groups had gaze and switch profiles which were compatible with an Elimination strategy. Early attention to distractors (e.g., shape) seems to be correlated with poor performance, whereas early alignments of learning items lead to better performance. Differences observed between strategies are discussed in an executive function framework.

Connecting eye movements with strategies.

The main focus is placed on the time course of gazes and switches during a trial. The important results are, first, that the fixation times for the learning items were longer than those for the generalization options at the beginning of the trials. Switches between the learning items were also very frequent at the beginning of each trial. This suggests that the learning items play an early anchoring role in novel noun generalization strategies. The importance of these early gazes towards the learning items is also shown by their predictive power, as indicated by the analyses of the initial gazes.

Second, gazes towards the generalization options and switches between learning items and generalization options appeared later in the trials, while there were far less switches between the available generalization options themselves. Thus, switches between learning and generalization options constituted the main pattern exhibited by children when focusing on a generalization solution.

One main aim was to test this data set against four possible generalization strategies, the “Elimination strategy”, the “Construction strategy”, the “Goal-Directed” strategy and the

“Salient Shape strategy”. Both the Elimination and the Construction strategies are compatible with an early focus (longer gazes and switches) on the learning items, which we interpret in terms of a search for and comparison of common features. The Goal-Directed strategy is less compatible with this early focus on the learning items because it predicts an early focus on the generalization items. An Elimination strategy calls for a large number of switches between learning items and available generalization options, which is what was observed. A Construction strategy would have predicted switches between available generalization options, that is comparisons of the generalization options with each other, which was not the case or, at least, much less so than the case of switches towards options.

The Goal-Directed strategy predicted an equal distribution of gazes between learning items and generalization options, or an early focus on generalization options (longer gazes towards generalization options than other items and more switches between generalization options than other switches) which, again, was not the dominant pattern. The pattern of answers is also less compatible with the Salient Shape hypothesis, which describes a bottom-up process driven by shape similarities. Indeed, this view predicts early gazes and switches towards the perceptually similar items, whereas our results show that the perceptual distractor was the object of shorter gazes than the learning items.

However, in the beginning of trials, the results also revealed more switches between learning items and the perceptual distractor than between learning items themselves and this might cast doubt on our interpretation. As predicted by the alignment view, perceptual distractors might attract early comparisons because of their salience. However, successful trials were characterized not by longer *gazes* at the perceptual distractor – which would have been a sign of a bottom-up process - but by longer gazes towards learning items. However, the switches between learning items and the perceptual distractor co-occurred with switches between the learning items and the thematically-related distractor and the taxonomic option.

These switches, which persisted until the taxonomic option was chosen over the other options, are compatible with a top-down decision process as the discussion of initial gazes will show.

Overall, the pattern of results is consistent with the response elimination strategy described for analogical reasoning [63, 64], in which items are aligned and eliminated until the best match is the only one left. It is also compatible with the “alignment-first” conception derived from the alignment hypothesis [53], which predicts that each option will be aligned until one is kept as the solution, either correct or incorrect.

Profiles and predictive power of the early gazes

Our fine-grained analysis of the children’s first five gazes revealed two important and related results. The first is that the ratio between learning-stimulus gazes and generalization-stimulus gazes (Learn/Gen) was predictive of taxonomic generalization performance up to the first three gazes in near generalization trials and up to the first four gazes in distant generalization trials. Hence, a high ratio of early gazes towards the learning items is predictive of a taxonomic choice. Again, this makes sense in terms of alignment and confirms the previous analyses run on gazes and switches. These early alignments between learning items are important in helping to build a deeper encoding of the learning items which, in turn, makes it possible to find the taxonomic option. This finding echoes those reported for analogical tasks that have also revealed the importance of early gazes towards the base domain items (A and B stimuli) before participants start to search for the task solution among the available generalization options [77, 78]. In both cases, it is interesting that children’s ability to control their search and fixate learning items not only improves with age, as was the case here, but is also positively correlated with success.

The second important and related result is that high-achieving children were able to selectively concentrate on the learning items at trial onset than low-achieving children. This

was confirmed by the higher Learn/Gen ratios for high-achieving children across the five gazes and supports the interpretation that achieving correct generalization may be dependent on children's ability to efficiently control their strategy by focusing on the learning items.

This result is analogous to results obtained in analogical reasoning tasks. As in the present experiment, the eye-movement profiles of the low-achieving participants in Thibaut and colleagues' studies on analogies (see also [79, 80]) were more biased towards the C item and the available options (C and the options defining the goal of the task) than those of high-achieving participants. Indeed, as shown by [81], when the analogy task required children to first analyze and identify the relation between A and B before they were allowed to see C and the options (C and D part of the analogy), the results were better than in a condition in which all the stimuli composing the analogy were shown simultaneously [82]. Thus, anchoring generalization in a thorough analysis of the learning items seems to be the first requirement before then comparing the learning items with the generalization items. In terms of strategy, this suggests that participants who follow a bottom-up strategy which requires them to attend to the shape (or other distractors) are less efficient. This is exactly what the analyses on the first gazes show. The same is true for those who follow the Task Goal-Directed strategy and distribute their attention across all the stimuli, including the generalization options, early on, at the risk of diluting task encoding by including all the stimuli rather than just the learning items. The latter hypothesis is interesting as it provides an interpretation in terms of processes (inhibition of the task's main goal) rather than in terms of content, that is in terms of what information is used (perceptual or deeper, or both). These two strategies are compatible with the profiles of children who directed more early gazes towards distractors (low achievers) and who might be those who failed to inhibit this information and were less successful in the task. In this respect, an analysis of six-year-old children's errors (paired t-test) revealed that they made significantly more perceptual errors than thematically related errors ($M_P = 38.58\%$, M_{Th}

= 11.35 %, $t(184) = 4.10$, $p < .001$, $d = 0.47$). This argues in favor of the interpretation that the children who made mistakes were the ones who chose the perceptual distractor instead of integrating all the provided information. These participants are those who followed the Salient-shape strategy.

Are generalization strategies mediated by age and task difficulty?

We manipulated age and conceptual distance as two factors known to influence generalization performance and which are thus potential mediators of generalization strategies. Behavioral data confirmed that both age and distance could play a role. Eye-tracking data revealed a main effect of age and of generalization distance, even though these two factors were not involved in any interaction with other factors. For example, data did not reveal a switch from elimination to construction strategies as task difficulty increased.

Age and generalization distance do, however, affect generalization performance. In the distant generalization condition (i.e., when the difficulty of the task increased), younger children made significantly more errors or had higher generalization RTs. In terms of strategies, there were main effects of age and generalization distance on the number of switches, with less switches for younger children and for distant generalization items. This might appear paradoxical since it might be expected that younger children would need to make more comparisons in order to find a solution and that distant items would require more comparisons. The latter result might suggest that younger children were unable to control the task until they had achieved a relevant encoding of the learning items and, indeed, they probably gave up attempting to do this earlier. This is consistent with their near-chance performance and the fact that they answered more rapidly in the difficult cases. Similarly, the more difficult, distant, generalization condition may also have elicited less comparisons overall (log of the number of switches: $M_{Near} = 0.46$, $M_{Distant} = 0.38$, $p < .001$, see results for

more details), most likely because children found it difficult to unify the items. Finally, the models of the predictive power of the Learn/Gen ratio for the initial gazes in the two generalization distance conditions showed that a longer series of gazes was more predictive in the distant generalization condition than in the near generalization condition (first four gazes were predictive versus first three gazes in the case of near generalization). Thus, it seems that it is not the number of gazes per se that matters for selecting the taxonomic match but rather the fact that they are targeted on learning items.

In sum, at any age, being a smart generalizer starts with systematic alignments of the learning items at the beginning of the trial, followed by systematic comparisons with the options later on. Comparing the options with one another, as would be predicted by a projection-first strategy or construction strategy (see above), does not seem to be the most important factor here, even though children did make these comparisons. Previous contributions on comparison designs have described novel noun learning as an alignment between learning items [37, 47]. This view is important but neglects the steps that occur after or in parallel with alignment, namely the steps involving comparisons with the options. Generalization also involves successive comparisons, as each option provides a set of possibilities for new alignment hypotheses. During generalization, children apply the vocabulary of dimensions that they have abstracted during learning, keep it in mind and resist irrelevant options. In other words, they start with learning items and then generalize the dimensions to the new items. The case of the classical single novel noun learning task is very different in this respect. Children use their lexical biases to target certain dimensions and find the stimuli that are similar along this dimension. It is more a “check and decide” process than the constructive process described here.

Limits of the present study

We focused on a small number of stimuli that we thought were representative of the issue. This has been the strategy followed by most developmental studies and seems a sensible one to replicate. In their study, Thibaut and French (2016) found that small subsets of the data led to the same strategies (near perfect correlation) as the full data set. Second, our experiment focused on object nouns. The extent to which our results can be generalized to other types of nouns, such as relational nouns [43], remains a matter of debate. Indeed, relational nouns involve at least two entities (e.g., being the “neighbor” of, “addition” of numbers). Thus, participants would have to analyze the two related entities and align the output of this comparison with the second set of stimuli that implement the same relational noun. Third, it would be interesting to study younger children. However, with the materials and the eye-tracker used here, we could not collect reliable data from younger children. Despite this, extending the age range remains a desirable objective.

Conclusion

The main purpose of the present study was to describe children’s generalization strategies and assess how these may be mediated by task difficulty. Overall, our results give a fairly clear picture of what a successful alignment looks like for children aged from five to eight years. Children mostly use what can be described as an elimination strategy, in which they first look longer at the learning items and make comparisons between them. Also, switches between learning items and available generalization options start very early in the task, as some are observed at the same time as the comparisons between learning items at the beginning of the trials. This dominant strategy was used in a robust way, whatever the age and task difficulty. Differences in generalization can therefore not be explained by strategy profile differences here. Two important aspects have been shown to determine the success of

conceptual generalization: the search pattern children use to navigate towards the solution and their early attention to learning items rather than available generalization options.

Acknowledgements

The first author benefited from a PhD scholarship from the Ministry of Higher Education and Research. The authors also thank the Agence Nationale de la Recherche (France) (ANR-18-CE28-0019-01 COMPARE, ANR Grant) and the Conseil Régional de Bourgogne Franche-Comté (PARI program) for their financial support, and Yannick Lagarrigue and members of the LEAD laboratory for helpful discussions.

Conflicts of interest: none

References

1. Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
2. Carey, S., & Bartlett, E. (1978). *Acquiring a single new word*.
3. Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
4. Baldwin, D. A. (1995). Understanding the link between joint attention and language. *Joint Attention: Its Origins and Role in Development*, 131, 158.
5. Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using Speakers' Referential Intentions to Model Early Cross-Situational Word Learning. *Psychological Science*, 20(5), 578–585. <https://doi.org/10.1111/j.1467-9280.2009.02335.x>
6. Gentner, D., & Namy, L. L. (2004). The Role of Comparison in Children's Early Word Learning. In *Weaving a lexicon* (pp. 533–568). MIT Press.
7. Gentner, D., & Namy, L. L. (2006). Analogical Processes in Language Learning. *Current Directions in Psychological Science*, 15(6), 297–301. <https://doi.org/10.1111/j.1467-8721.2006.00456.x>
8. Smith, L. B. (1999). Children's noun learning: How general learning processes make specialized learning mechanisms. *The Emergence of Language*, 277–303.
9. Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–1568. <https://doi.org/10.1016/j.cognition.2007.06.010>
10. Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, 10(2), 201–224. [https://doi.org/10.1016/0885-2014\(95\)90009-8](https://doi.org/10.1016/0885-2014(95)90009-8)
11. Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, 114(2), 245–272. <https://doi.org/10.1037/0033-295X.114.2.245>
12. Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, 8(2), 113–139.

13. Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*.
Mit Press.
14. Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. *Child Development, 61*(5), 1461–1473.
15. Diesendruck, G., Gelman, S. A., & Lebowitz, K. (1998). Conceptual and linguistic biases in children's word learning. *Developmental Psychology, 34*(5), 823.
16. Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development, 74*(1), 168–178.
17. Imai, M., & Haryu, E. (2004). The nature of word learning biases and their roles for lexical development. *Weaving the Lexicon*, 411–444.
18. Kucker, S. C., Samuelson, L., Perry, L., Yoshida, H., Colunga, E., Lorenz, M., & Smith, L. (2019). Reproducibility and a Unifying Explanation: Lessons from the Shape Bias. *Infant Behavior & Development, 54*, 156–165. <https://doi.org/10.1016/j.infbeh.2018.09.011>
19. Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321.
20. Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological science, 13*(1), 13-19.
21. Kucker, S. C., McMurray, B., & Samuelson, L. K. (2018). Too much of a good thing : How novelty biases and vocabulary influence known and novel referent selection in 18-month-old children and associative learning models. *Cognitive science, 42*, 463-493.
22. Sloutsky, V. M. (2010). From perceptual categories to concepts : What develops? *Cognitive science, 34*(7), 1244-1286.

23. Sloutsky, V. M., & Fisher, A. V. (2004). Induction and Categorization in Young Children : A Similarity-Based Model. *Journal of Experimental Psychology: General*, 133(2), 166-188. <https://doi.org/10.1037/0096-3445.133.2.166>
24. Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children : A dumb attentional mechanism? *Cognition*, 60(2), 143-171.
25. Smith, L., & Samuelson, L. (2006). An Attentional Learning Account of the Shape Bias : Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental psychology*, 42, 1339-1343. <https://doi.org/10.1037/0012-1649.42.6.1339>
26. Madole, K. L., & Oakes, L. M. (1999). Making Sense of Infant Categorization : Stable Processes and Changing Representations. *Developmental Review*, 19(2), 263-296. <https://doi.org/10.1006/drev.1998.0481>
27. Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
28. Gelman, S. A., & Meyer, M. (2011). Child categorization. *Wiley interdisciplinary reviews. Cognitive science*, 2(1), 95-105. <https://doi.org/10.1002/wcs.96>
29. Markman, A., & Ross, B. H. (2003). Category use and category learning. *Psychological bulletin*, 129(4), 592.
30. Waxman, S. R. (1990). Linguistic biases and the establishment of conceptual hierarchies : Evidence from preschool children. *Cognitive Development*, 5(2), 123-150. [https://doi.org/10.1016/0885-2014\(90\)90023-M](https://doi.org/10.1016/0885-2014(90)90023-M)
31. Waxman, S. R., & Hatch, T. (1992). Beyond the basics : Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language*, 19(1), 153-166
32. Gentner, D., & Namy, L. L. (1999). Comparison in the Development of Categories. *Cognitive Development*, 14(4), 487–513. [https://doi.org/10.1016/S0885-2014\(99\)00016-7](https://doi.org/10.1016/S0885-2014(99)00016-7)

33. Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*(1), 1–27.

[https://doi.org/10.1016/0010-0285\(84\)90002-1](https://doi.org/10.1016/0010-0285(84)90002-1)

34. Samuelson, L. K., Schutte, A. R., & Horst, J. S. (2009). The dynamic nature of knowledge: Insights from a dynamic field model of children's novel noun generalization. *Cognition*, *110*(3), 322-345. <https://doi.org/10.1016/j.cognition.2008.10.017>

35. Horst, J. S., & Twomey, K. E. (2013). It's Taking Shape: Shared Object Features Influence Novel Noun Generalizations. *Infant and Child Development*, *22*(1), 24-43.

<https://doi.org/10.1002/icd.1768>

36. Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference? *Cognition*, *112*(1), 105-119.

<https://doi.org/10.1016/j.cognition.2009.03.012>

37. Augier, L., & Thibaut, J.-P. (2013). The benefits and costs of comparisons in a novel object categorization task: Interactions with development. *Psychonomic Bulletin & Review*, *20*(6), 1126–1132.

38. Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology*, *107*(3), 280–290. <https://doi.org/10.1016/j.jecp.2010.04.017>

39. Thibaut, J.-P., & Witt, A. (2023). Children's generalization of novel names in comparison settings : The role of semantic distance during learning and at test. *Journal of Experimental Child Psychology*, *234*, 105704. <https://doi.org/10.1016/j.jecp.2023.105704>

40. Thibaut, J.-P. (1995). The abstraction of relevant features by children and adults: The case of visual stimuli. *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society*, *17*, 194–199.

41. Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, *36*(5), 571–581.
42. Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language*, *36*(1), 201–224.
43. Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children’s learning of relational categories. *Child Dev*, 1173–1188.
44. Thibaut, J. P., & Witt, A. (2015). Young children’s learning of relational categories: Multiple comparisons and their cognitive constraints. *Frontiers in Psychology*, *6*.
<https://doi.org/10.3389/fpsyg.2015.00643>
45. Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons: A Meta-Analytic Review. *Educational Psychologist*, *48*(2), 87–113.
<https://doi.org/10.1080/00461520.2013.775712>
46. Childers, J. B. (Ed.). (2020). *Language and Concept Acquisition from Infancy Through Childhood: Learning from Multiple Exemplars*. Springer International Publishing.
<https://doi.org/10.1007/978-3-030-35594-4>
47. Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children’s categorization. 14.
48. Hammer, R., Bar-Hillel, A., Hertz, T., Weinshall, D., & Hochstein, S. (2008). Comparison processes in category learning: From theory to behavior. *Brain Research*, *1225*, 102–118.
<https://doi.org/10.1016/j.brainres.2008.04.079>
49. Thibaut, J.-P., & Witt, A. (2017). Generalizing novel names in comparison settings: Role of conceptual distance during learning and at test. *CogSci 2017 : 39th Annual Meeting of the Cognitive Science Society*, 3314–3319. <https://hal-univ-bourgogne.archives-ouvertes.fr/hal-01573610>

50. Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy*. *Cognitive Science*, 7(2), 155–170. https://doi.org/10.1207/s15516709cog0702_3
51. Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, 131(1), 5.
52. Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, 25(4), 524–575.
53. Markman, A. B., & Gentner, D. (1993a). Splitting the Differences: A Structural Alignment View of Similarity. *Journal of Memory and Language*, 32(4), 517–535. <https://doi.org/10.1006/jmla.1993.1027>
54. Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything : Appearance similarity versus kind information in early induction. *Psychological Science*, 18(2), 179- 185.
55. Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, 13(6), 258- 263. <https://doi.org/10.1016/j.tics.2009.03.006>
56. Waxman, S. R., & Gelman, S. A. (2010). Different kinds of concepts and different kinds of words : What words do for human cognition. *The making of human concepts*, 101- 130.
57. Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
58. Rayner, K., Shen, D., Bai, X., & Yan, G. (2018). *Cognitive and cultural influences on eye movements*. CRC Press.
59. Soluch, P., & Tarnowski, A. (2013). Eye-tracking methods and measures. *Translation Studies and Eye-Tracking Analysis*. Frankfurt Am Main: Peter Lang, 85–104.
60. Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52, 45–56.

61. Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*(1), 1–38.
62. Holyoak, K. J. (2012). Analogy and Relational Reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (1st ed., pp. 234–259). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199734689.013.0013>
63. Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984b). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence*, *8*(3), 205–238. [https://doi.org/10.1016/0160-2896\(84\)90009-6](https://doi.org/10.1016/0160-2896(84)90009-6)
64. Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984a). Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence*, *8*(3), 205–238. [https://doi.org/10.1016/0160-2896\(84\)90009-6](https://doi.org/10.1016/0160-2896(84)90009-6)
65. Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review*, *104*(3), 427.
66. Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*(2), 220.
67. Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, *27*(2), 195–232. [https://doi.org/10.1016/0022-0965\(79\)90044-4](https://doi.org/10.1016/0022-0965(79)90044-4)
68. Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development*, *38*, 10–26.
69. Rosch, E., & Turbiaux, M. (1976). Classifications d'objets du monde réel : origines et représentations dans la cognition. *Bulletin de psychologie*, *29*(325), 242-250.
70. Waxman, S. R., & Hatch, T. (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language*, *19*(1), 153–166.

71. Duchowski, A. (2007). Eye tracking techniques. In *Eye tracking methodology* (pp. 51–59). Springer.
72. Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, 8(3), 259–272.
73. New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36(3), 516–524. <https://doi.org/10.3758/BF03195598>
74. Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, 4(1), 6–14. [https://doi.org/10.1016/S1364-6613\(99\)01418-7](https://doi.org/10.1016/S1364-6613(99)01418-7)
75. Duchowski, A. J. (2007). *Eye-tracking methodology: theory and practice*. London: Springer-Verlag
76. French, R. M., Glady, Y., & Thibaut, J.-P. (2017). An evaluation of scanpath-comparison and machine-learning classification algorithms used to study the dynamics of analogy making. *Behavior Research Methods*, 49(4), 1291–1302.
77. Guarino, K. F., Wakefield, E. M., Morrison, R. G., & Richland, L. E. (2022). Why do children struggle on analogical reasoning tasks? Considering the role of problem format by measuring visual attention. *Acta Psychologica*, 224, 103505.
78. Thibaut, J.-P., French, R., & Vezneva, M. (2010). The development of analogy making in children: Cognitive load and executive functions. *Journal of Experimental Child Psychology*, 106(1), 1–19. <https://doi.org/10.1016/j.jecp.2010.01.001>
79. Starr, A., Vendetti, M. S., & Bunge, S. A. (2018). Eye movements provide insight into individual differences in children’s analogical reasoning strategies. *Acta Psychologica*, 186, 18–26. <https://doi.org/10.1016/j.actpsy.2018.04.002>

80. Vendetti, M. S., Starr, A., Johnson, E. L., Modavi, K., & Bunge, S. A. (2017). Eye Movements Reveal Optimal Strategies for Analogical Reasoning. *Frontiers in Psychology*, 8.

<https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00932>

81. Glady, Y., French, R. M., & Thibaut, J.-P. (2016). *Comparing competing views of analogy making using eye-tracking technology*.

82. Green, A. E. (2016). Creativity, Within Reason: Semantic Distance and Dynamic State

Creativity in Relational Thinking and Reasoning. *Current Directions in Psychological*

Science, 25(1), 28–35. <https://doi.org/10.1177/0963721415618485>

Supporting information

S1 Table. Full list of materials for the Close vs. Far learning and the Near vs. Distant generalization conditions.

S2 Table. Perceptual conceptual and semantic similarity ratings by category

S3 Table. Results of Spearman correlations between ratio Learn/Gen indicating children's early attention to learning items rather than available generalization options and their percentage of taxonomic answers.

S4 Table. Details of the ANOVA run on M3 model including Learn/Gen ratio for the first 3 gazes, in near generalization settings. Controlling for age and learning distance.

S5 Table. Details of the ANOVA run on M5 model including Learn/Gen ratio for the first 5 gazes, in distant generalization settings. Controlling for age and learning distance.

S6 Data. Full data set

Chapter 2 - Relational noun generalization strategies: A developmental eye-tracking analysis.

The following chapter contains the article by E. Stansbury, A. Witt, P. Bard, J-P. Thibaut, (in preparation) Relational noun generalization: a developmental eye-tracking analysis.

Abstract

Comparison and novel word generalization tasks are widely recognized as mapping tasks, while little is known of the search strategies that participants in such tasks implement. In the present experiment, we address this issue by analyzing eye movements. In a comparison setting, we tested 5-year-old children's and adults' relational noun generalizations while manipulating the semantic distance between learning items and between learning and generalization items. The eye movement data revealed that both adults and children follow a projection-first strategy that predicts attention to learning items to extract a representation of the targeted relation before projecting the relation in the generalization domain to find a solution. Interestingly, the data also revealed that the differences between children's and adults' strategies were not in the type of strategy but in the efficiency of the strategy's implementation. We shall discuss this result in an executive framework and consider its implications for the mental representations used by participants.

Keywords:

Comparison, Relational noun, Eye-tracking, Search strategies, Novel noun extension, Generalization, Semantic distance, Concepts

Introduction

Children remarkably succeed in learning words with limited learning stimuli (Bloom, 2002; Carey & Bartlett, 1978; Murphy, 2002). However, their first attempts at generalizing words are frequently biased. Among the numerous explanatory theories of novel word generalization (Baldwin, 1995; Gentner & Namy, 2004a; L. B. Smith, 1999; Tomasello & Akhtar, 1995; Xu & Tenenbaum, 2007b), a prominent view is that children's generalization is biased towards particular dimensions of the learning stimuli (Jones & Smith, 1993; Markman, 1989; Waxman & Kosowski, 1990). Generalization biases arise when children use one of the items' dimensions preferentially to generalize the novel word. For example, in shape bias, children generalize object names to items with similar shapes (Kucker et al., 2019b; e.g., Landau et al., 1988c) rather than items sharing deeper conceptual similarities. In the case of relational nouns generalization, participants are asked to generalize a word designating a relation between two items to a novel pair of items related in the same way. For example, the

relation “cutter for” unites a knife and an orange, the knife being the cutter for the orange, and could be generalized to the pair “knife and watermelon.” In these relational noun generalization tasks, a frequent bias is the taxonomic bias that leads children to generalize relational nouns to pairs of items from the same taxonomic category (e.g., two knives or two oranges) rather than items that share the correct relation (Gentner & Toupin, 1986; Richland et al., 2006).

In the present experiment, we study relational noun generalization, which tends to be more difficult for children because relational nouns do not refer to intrinsic, perceptually stable, and easily identifiable properties. A relational generalization task will be adapted here, which involves showing participants items sharing the targeted relation (i.e., “cutter for,” a knife and an orange, the knife is the operator that cuts and the orange the entity that is cut) and asking them to find an operator item among available options that share the same relation with a given entity target.

Constraining word learning by comparison settings

Research has recently focused on learning designs that may constrain children’s learning of word meaning and favor taxonomic generalization. For example, Gentner et al. (2011) tested which conditions would lead to better generalizing relational nouns. They introduced one (Experiment 1) or several (Experiments 2 and 3) learning pairs built around two familiar objects connected by a familiar relation (e.g., “cutter for”), one being the operator (e.g., a knife), the other the entity (e.g., an orange). Items were presented as X is the “dax” for Y (e.g., the knife is the “dax” for the orange). At test, in all the experiments, a target entity (e.g., a sheet of paper) was introduced with three alternatives (i.e., a relational match— a pair of scissors—, a taxonomic match— a pile of sheets of paper—, and a thematic match— a pencil—). Children were asked to show which stimulus among the alternatives was the “dax” for the piece of paper (Relational Label condition). Authors found that more relational matches were made when children had compared multiple items at learning rather than when they had learned in no-comparison settings.

It is generally considered that this type of novel relational noun generalization paradigm involves alignments between task stimuli and conceptual domains (Gentner et al., 1993; Gentner & Markman, 1997; Gick & Holyoak, 1983; Holyoak, 2012; A. B. Markman & Gentner, 1993a, 1993b, 1996; Medin et al., 1993). From this point of view, comparison at learning benefits generalization because it helps participants align stimuli’s shared properties,

progressively neglects irrelevant perceptual properties, and considers non-obvious deeper commonalities. A common representation is extracted from this mapping process that is considered to be based on dimensions that are more pertinent for future generalizations than a representation extracted from a no-comparison learning setting.

Extensive research has confirmed and affirmed our understanding of settings that constrain learning and favor word generalization. However, little is still known about the solving strategies used to process comparison settings and generalize novel words or the steps that lead to correct generalization. Therefore, the question we address here is: *How* does one come to a generalization decision? Are specific strategies followed to reach a generalization decision? If such strategies exist, understanding their temporal dynamics, their development in individuals, or how their implementation impacts successful generalization are important questions for a better understanding of word generalization.

We will address these questions with an eye movement analysis of a relational noun generalization task with 5-year-old children and adults to catch early and developed relational noun strategies and compare their profiles.

Investigating strategies in mapping tasks

To our knowledge, strategies followed to generalize a relational noun in a comparison learning design have not yet been explored. We aim to uncover these strategies using eye movements because they are considered a precise investigation method of participants' attention to visual stimuli (Soluch & Tarnowski, 2013). Such data also reflects interactions between cognitive processes and external visual stimuli (Kahneman, 2011; Rayner et al., 2018).

Although this is a new way to consider novel relational noun generalization tasks and to investigate the mapping processes and the strategies temporal dynamics behind the behavioral data already known, it must be mentioned that this approach has already been carried out in analogical reasoning tasks (Thibaut & French, 2016; Starr et al., 2018), verb generalization tasks (Childrens et al., 2019) another relational word generalization task, and novel noun generalization tasks (Stansbury et al., submitted). Indeed, Thibaut and French (2016) analyzed eye-tracking data in an analogical task, another comparison-generalization task, and showed that children's exploration profiles differed from adults' profiles and that errors and correct answers did not display the same search-strategy profile. Other authors also revealed analogical reasoning strategies in adults by analyzing their gaze sequences (Vendetti

et al., 2017b). Interestingly, they also showed that the most used and optimal strategy was a projection-first strategy in which participants explored the learning pair before generalizing the analogical relation in the target domain. Starr and colleagues (2018) investigated children's analogical reasoning strategies. They showed that children tended to focus on the task's goal rather than the learning stimuli (Starr et al., 2018b). However, significant differences existed between children, and those who adopted a strategy involving paying attention to the learning stimuli before turning to the task's goal were more successful.

We suggest that comparison-analogical reasoning tasks and comparison-relational noun generalization tasks can involve similar mapping processes. Indeed, in both these tasks, participants must consider pairs of items in a source domain and compare these pairs to extract a representation of their common relation. They also have to analyze the options in a target domain and extend the regularity/ies extracted from the source domain to find the task's solution. In typical analogical reasoning tasks, participants investigate the A-B pair in the source domain and search for D in the target domain so that D is to a given item C what B is to A. In the case of relational noun generalization, participants investigate relational entity-operator pairs in the source domain and search among available options for an item R (the relational operator) to match a given target entity so that the relation between R and the target entity matches the relation between the learning entity-operator pairs. Both these tasks involve mapping items in two domains: a source domain for collecting information and a target domain in which to find a solution to match a given item. In this sense, parallels can be made between mapping strategies in analogical reasoning and relational noun generalization.

Before we come to hypotheses regarding possible relation noun generalization strategies, we shall therefore summarize mapping strategies described in analogical reasoning. These strategies will then be used as references to hypothesize equivalent strategies for relational noun generalization.

Strategies in analogical reasoning tasks

Several strategies have been described in analogical reasoning. One strategy has been called the *Alignment-first hypothesis* and involves comparing items in the target domain to equivalent items in the source domain, eliminating inadequate target items until only one is left, which will be selected as the answer (Markman & Gentner, 1996)

A second strategy is the *Projection-first hypothesis*, which involves first studying the source pair (A and B) and extracting the relation unifying the pair. Second, applying this

relation in the target domain, item D is found to match C as B matches A. (Hummel & Holyoak, 1997, 2003; Sternberg & Rifkin, 1979).

A third strategy has been described by Thibaut and French (2016), the *Task goal-directed strategy*. In their analogy paradigm, younger children had more gazes toward the solution set from the beginning than older children and adults. The authors interpreted this pattern in terms of children's difficulty inhibiting the task's main goal: "finding the one that goes with C." In our lexical comparison-generalization design, the main task is also to select one item among the solution set and could also a priori give this type of search profile.

The present study. Search strategies in relational noun generalization tasks

Considering the similarities between analogical reasoning tasks and relational noun generalization tasks, it is reasonable to suggest that relational noun generalization strategies (i.e., mappings during the relational noun generalization strategy) may display similar profiles to analogical reasoning hypothesized strategies. In the present study, we investigated these strategies in a comparison design.

Our first aim is to describe these strategies' temporal dynamics and interpret them in the light of the strategies hypothesized in analogical reasoning. Relational noun strategies will be captured by the temporal organization of participants' search in terms of gaze length (a measure of the depth of analysis of a stimulus), number of switches between stimuli (a measure of which stimuli are compared), and early gaze location (a measure of early attention orientation). We will focus on the proportion of fixation times which reveal participants' gaze distribution, hence which items they have gazed at the most and consider the best source of information; switches between AOI, which reveals a comparison between items and participant's search actions; and first gaze location which reveal early attention and search sequences.

We predict that participants' relational noun generalization strategies will follow an alignment or a projection strategy, or a task goal-directed strategy. This, however, does not exclude the possibility that relational generalization strategies may have profiles that can combine aspects from more than one of these strategies. We suggest that, in relational noun generalization, an alignment strategy (i.e., aligning equivalent items to find a solution) will first involve comparing learning entities across learning pairs (i.e., entities from different learning pairs) and learning operators across learning pairs (i.e., operators from different learning pairs), before moving on to search for a solution by comparing learning entities with

the tasks target entity and learning operators with each available choice. We suggest that, in relational noun generalization, a projection strategy (i.e., projecting a relation extracted in the learning domain into the target domain) will involve considering each learning pair as a whole (i.e., entity and operator) to extract a mental representation of their relation, before matching the target entity to available options until the correct match is found and chosen (i.e., the solution that fits the mental representation the best). A task goal-directed strategy would be characterized by early gazes at the tasks' generalization items (i.e., the entity for which a match must be found and the available choices). This strategy does not give strict predictions about children's following mappings that may thus be better described by one of the other strategies.

A question that arises from generalization strategy investigation is what may affect search strategies. This is a vast question that could be investigated in itself. However, our aim in the present study is only to consider the effect on generalization strategies of a factor well known to affect relation known generalization performances: semantic distance between items (Thibaut & Witt, 2015). We manipulated the semantic distance between learning pairs and between learning pairs and generalization options. The reasoning is straightforward: the semantic distance between items affects participants' generalization performance (Thibaut & Witt, 2015), and the explanation of this effect is open: it may be that the conceptual distance at learning and generalization affects the mental representations built (Hummel & Holyoak, 1997) or the task's cognitive load (Thibaut et al., 2010, for analogical reasoning; Thibaut & Witt, 2017, for relational noun generalization). It is interesting to keep in mind that these two hypotheses may actually be two parts of the same process; the representation built at learning may directly impact the cognitive load of the search for a generalization solution. We suggest that semantic distance learning items and between learning items and generalization items, that has been shown to affect generalization performances, also affects generalization strategies. Suppose a broader semantic distance between learning items, for example, induces participants to build a representation with broader categorical boundaries. In that case, participants may be more sensitive to distractors and have a less systematic strategy towards the correct relational response. As participants' search is directed towards distractors, they must engage in more flexible switches and inhibition away from them, which would entail more cognitive cost. Another simpler option is that learning from distant learning items may in itself be cognitively more taxing, and fewer cognitive resources may be available for generalization search. In the same way, searching with generalization items more distant from

learning items may in itself be a more complex process than searching with nearer generalization items. Discovering an effect of semantic distance, or its absence, will help discuss at what level semantic distance between items affects generalization.

The last purpose is to investigate how strategies may develop and how their implementation may impact successful generalization. If any exist, developmental differences will be found by comparing 5-year-old children's and adults' strategies. Differences in strategy implementation, if any exist, will be found by comparing trials that led to different outcomes. We present these questions together because it is likely that development and successful strategy implantation are related. Thibaut and French (2016) have shown, in an analogical reasoning task, that the distribution of gazes in the first third of trials could predict whether the trial's outcome would be a correct generalization or an error. This implies that trials with different outcomes may be characterized by strategies with different patterns from the trial's onset. Hence, we shall pay special attention to early strategy sequences described by the first gazes' location and early fixations and switches. We predict that more successful generalization strategies (i.e., adults' strategies over children's or strategies from correct trials over errors) will be characterized by earlier and more attention to learning items than items in the target domain.

Method

Participants

Seventy-six 5-year-old children (mean age = 5.45 years; range: 4.6 to 6.1 years) were tested in a quiet room at their school. The children's parents received a written explanation about the research project and returned written informed consent.

Thirty-one adults (mean age = 18.5 years; range: 17 to 23 years) were recruited. They were all undergraduate psychology students from the University of Burgundy participating for points for their grades. They were tested at the laboratory after reading the research project's description and giving written informed consent.

All participants were randomly assigned to one of the two experimental conditions (close comparison, 37 children and 17 adults, or far comparison, 39 children and 14 adults).

The procedure followed the Declaration of Helsinki (1964). Written informed consent was obtained from adults and children's guardians; all data was anonymized, and all participants were informed of their right to withdraw at any time. The study was carried out

under the official agreement (convention n°: 2019-0679 and endorsement n°: 2020-0566) between the Academic Inspection of the French National Education Ministry, the University of Burgundy (“Inspection Académique de Côte d’Or”), and our laboratory.

Materials

Fourteen experimental sets of objects from Thibaut & Witt (2017) were used, and four sets were built for the warm-up trials. Each set was associated with a relational category (e.g., *cutter for*, *baby of*, *travel space for*, *food product of*, etc.) and contained 14 pictures of objects, either for the learning or the generalization phase.

Learning stimuli were learning pairs composed of an entity (E) and an operator (O) (e.g., for a trial, such as “a knife is a cutter for watermelon,” the watermelon is the entity, and the knife is the operator). As mentioned, we manipulated the distance between learning pairs (close, far). The close learning trials were built with a standard learning pair (L) and a close learning pair (Lc) conceptually close to the standard learning pair (e.g., kitchen knife and orange). The far learning trials were built with the standard pair (L) and a far leaning pair (Lf) conceptually less close to the standard learning pair than the close learning pair (see Figure 1).

Generalization phase items were sets of four pictures. -A set had a TargetEntity (TE) and three choices each related to the TargetEntity: a Relational match (R), a Taxonomic match (Ta), and a Thematically related match (Th) (e.g., a paper as the TargetEntity, a pair of scissors as the relational match, a paper as the taxonomic match, a pencil as the theme-related match). We manipulated generalization distance (i.e., the conceptual distance between the generalization relational pair and the standard learning pair) by building two different sets for each category, a near and a distant set. In the near set (e.g., paper, scissors, pencil), the TargetEntity was conceptually nearer to the standard learning entity than the TargetEntity from the distant generalization set (e.g., beard, shaver, toothbrush).

Each participant saw 14 trials, all in the same learning condition, either close or far, with 7 trials in the near generalization condition and 7 in the distant generalization condition.

Independent similarity ratings from 61 students were collected to control the material (i.e., semantic distances, relations between TargetEntity and available choices). Participants judged the similarity between entities and between operators in standard and close pairs and standard and distant pairs. The close pairs’ items ($M = 6.15$, 95%CI [5.98, 6.31]) were rated as significantly closer to the standard pairs’ items than the far pairs’ items ($M = 4.98$, 95%CI [4.81, 5.16]), $t(26) = 4.47$, $p < .001$. In addition, the entities and operators of the close pairs

($M_{Entities} = 5.72$, 95% CI [5.57, 5.87]; $M_{Operators} = 5.78$, 95% CI [5.58, 5.98]) were found to be more similar than those of the far pairs ($M_{Entities} = 3.82$, 95% CI [3.52, 4.12]; $M_{Operators} = 4.49$, 95% CI [4.27, 4.72]), respectively $t(13) = 6.16$, $p > .0001$, and $t(13) = 6.43$, $p < .001$.

Participants also rated semantic similarity between learning and relational pairs in generalization sets (i.e., Target entity and relational match) and between entities or operators from learning and generalization pairs (i.e., learning entity-TargetEntity or learning operator-relational match). Near generalization pairs (e.g., paper, scissors) ($M = 4.51$, 95% CI [4.33, 4.69]) were judged to be semantically more similar to learning pairs than distant generalization pairs (e.g., beard, shaver) ($M = 3.59$, 95% CI [3.31, 3.86]), $t(26) = 2.64$, $p < .02$.) Entities or operators in the near generalization sets (e.g., paper and scissors, respectively) were more similar to the entities (e.g., watermelon, orange and meat) or operators (e.g., knife1, knife2 and cleaver) in the close learning pairs than were those from distant generalization sets (e.g., bearded face and shaver, respectively). Entities in the near generalization sets were significantly more similar to entities in the learning pairs ($M = 2.90$, 95% CI [2.55, 3.26]) than entities in the distant generalization sets ($M = 1.65$, 95% CI [1.33, 1.96]), $t(13) = 3.98$, $p = .001$ ($< .0083$, Bonferroni-corrected p-value threshold), and operators in the near generalization sets were significantly more similar to operators in the learning pairs ($M = 3.61$, 95% CI [3.39, 3.83]) than operators in the distant generalization sets ($M = 2.74$, 95% CI [2.50, 2.98]), $t(13) = 3.99$, $p = .001$)

Fourteen different bisyllabic labels (pseudo-words) [62], easier to remember than monosyllabic pseudo-words, were used to name the learning objects in each trial. Syllables were of the CV type, which is the dominant word structure in French (from Lexique.org, bari, buxi, daxo, jito, malan, missi, muno, nati, peco, pina, rula, sefu, soki, vira, youma, zatu [63])

We used photos (500x500 pixels) displayed on a Tobii T120 eye-tracker device. The experiment was run with E-prime[®] software.

Precautions were taken to reduce the impact of the task's design on eye movements. Four different picture layouts on the screen were used to counterbalance any picture presentation effect. Learning pairs were presented vertically (one above the other) or horizontally (next to each other). In both cases, pairs were presented either left to right or right to left (i.e., layout 1, vertical-left to right; layout 2, vertical-right to left; layout 3, horizontal-left to right; layout 4, horizontal-right to left). All stimuli appeared on the screen at the same time. Training instructions and trials were used to explain the task so that the experimenter could avoid using hand movements during the test trials. Colored picture

boundaries and a fine black line in the center of the screen were used to help children orientate themselves and reduce their need for experimenter hand movements.

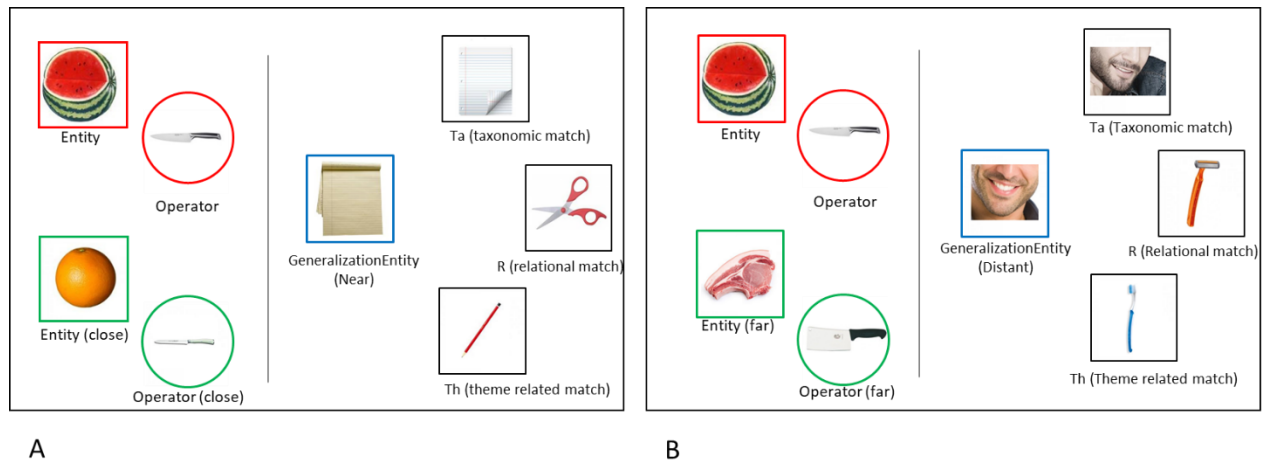


Figure 1: Example of a set of stimuli adapted for the relation “cutter for” crossing learning distance (close, far) and generalization distance (near, distant).

Note: Item names do not appear in the experiment. A crosses close learning and near generalization; B crosses far learning and distant generalization. Examples are in layout 1.

Procedure

Task presentation and warm-up: When the participants were settled in front of the computer with the experimenter, the task was presented as a game with a puppet, Yoshi. The puppet was used to make the task more attractive for children and to frame the use of non-existing names. Adults were advised that the task had been built for children and that they would see the task in the same way as children for experimental reasons. The following instructions were given: *"Hello, we are going to play a game with Yoshi. Yoshi lives in a faraway country and speaks a special language. In this game, we are going to play with words Yoshi uses for things. We are going to help him to sort some objects."* Participants then saw four practice trials during which the experimenter first showed the items as they gave the instruction and progressively stopped hand movements.

Eye-tracking calibration: The eye-tracking device was calibrated using the Tobii calibration procedure. Two calibrations were done, one between the warm-up and the first test trials and another between the 7th and 8th test trials. The second calibration was done to avoid losing data. Young children often move their heads during the task, leading the eye-tracker to lose track of their eye movements. A calibration mid-task enabled the eye-tracker to find eyes that had been lost and continue recording data.

Test: For each trial, all stimuli appeared on the screen and stayed in view until the end of the trial. The experimenter gave the following instruction (example in the case of material illustrating the relation “*cutter for*” in the close learning - near generalization condition): “*The big knife is the buxy for the watermelon. And the smaller knife is the buxy for the orange*”. The experimenter continued: “*But the paper does not have a buxy. We need to find the buxy for the paper. Can you help Yoshi and show him which object is the buxy for the paper?*” This procedure was repeated fourteen times. Items of a trial disappeared when the experimenter selected the items the participant chose. Trials were separated by a slide with a central fixation cross for 3 seconds. The presentation order and the items’ positions on the screen were automatically randomly assigned. Labels were interchanged among pairs across participants.

Results

Behavioral data

Performance was measured by the percentage of relational generalizations (i.e., percentage of relational choices) and children’s reaction times in seconds.

Generalization scores

We first analyzed percentages of relational choices. We ran a three-way repeated measures ANOVA on participants’ percentage of correct answers with age (child, adult) and learning distance (close, far) as between factors and generalization distance (near, distant) as a within factor. The analysis revealed a main effect of age, $F(1,103) = 91.94, p < .001, \eta^2_p = .47$. Children make less relational generalizations than adults ($M_{child} = 60.4, M_{adult} = 96.1$). One sample T-tests confirm that in all conditions, children scored above chance. ($p < .001$)

Reaction times

We ran a three-way repeated measures ANOVA on participants’ reaction times for correct answers with age (child, adult) and learning distance (close, far) as between factors and generalization distance (near, distant) as a within factor. The analysis revealed a main effect of age, $F(1,87) = 81.13, p < .001, \eta^2_p = .48$, adults answer faster than children ($M_{child} = 12.47, M_{adult} = 6.85$).

Taken together these results show that 5-year-old children’s answers are not affected by the conceptual distance between the task’s items, unlike what was hypothesized. A

possible interpretation of this result, considering that the children's percentage of correct answers is relatively high, is that the task is easy for 5-year-old children. Although the task is easy for 5-year-old children, the age-related performance suggests that the task is appropriately designed to investigate the potential effect of age on generalization performance and related strategies.

Eye-tracking data: Strategy Description

Conceptual distances

Learning and generalization distance did not affect the percentage of relational answers or reaction times (cf. previous section). For this reason, it was possible that conceptual distance between items may have no or little effect on relational noun generalization strategies.

We tested learning and generalization distances' effect on eye-tracking indexes (i.e., the proportion of fixation times and log of the number of switches between items) by running repeated measures ANOVAs on each index with age group (child, adult) and learning distance as between-subject factors and time slice (beginning, middle, end) and switch or AOI type as within-subject factors. The results of interest for seeing if conceptual distance effect's strategies are the interactions between distance (i.e., learning or generalization distance factor) and eye-tracking index patterns (i.e., AOI or switch type factor) over time (i.e., time slice factor)

Learning distance did not interact with switch type over time, $F(7,721) = 0.26$, $p = 0.97$, $\eta_p^2 = .003$). Learning distance interacted with AOI type over time $F(10,890) = 1.91$, $p = 0.041$, $\eta_p^2 = .02$). This effects size is very small, and post hoc analysis revealed that there were no differences between eye movement patterns in close and far learning cases. Considering the small effect of learning distance, we excluded this factor from further analyses to simplify the design and improve the result's robustness.

Generalization distance did affect eye movement patterns and will be described and discussed below.

Eye-tracking data collection and processing

Data from three children was removed because eye-tracking calibration was unsuccessful. Inadequate contrast between eye color, room lighting, and screen lighting made the pupil too difficult to detect.

Missing data arose when eyes were lost episodically during the task, resulting in the absence of eye position recording for a short period. We excluded trials for which more than 50% of eye positions had been lost, and we excluded subjects for whom more than 50% of trials had been excluded. We lost a total of ten subjects completely. The remaining participants lost, on average, 1.5 trials.

As in Thibaut and French (2016), eye-tracking data was divided into three equal time slices, that is, the beginning, the middle, and the end of a trial, in order to capture the time dynamics of a trial. A time slice's length was equal to a third of a trial's total fixation time. We first focus on AOIs and Switches (transitions). Gaze duration (or looking time) for areas of interest (i.e., AOIs towards the items themselves) tells which and for how long stimuli are attended to while solving the problem and reveals the depth of the item's processing. Switches between items (or saccades or transitions) tell us which items are compared and can be interpreted as an attempt to find a relation between X and Y (see Duchowski, 2007). The two measures reveal different aspects of the search, as a participant can study an item for a long time without comparing it with other items, and gaze duration, unlike transitions, tells us nothing about which items are compared during trials.

We applied normal transformations to data when necessary to fit the ANOVA normality requirement. Log transformations were applied to the number of switches. We applied Greenhouse-Geisser or Huynh-Feldt corrections to F scores when sphericity scores required them.

Fixation times

We analyzed children's proportion of fixation times during trials. Following standard practice (e.g., Liversedge & Findlay, 2000), a gaze to an AOI was considered a fixation when both eye positions were recorded inside the AOI's boundaries for longer than 60 ms. Proportions of fixation times met normal distribution ANOVA requirements.

We ran a four-way repeated measure ANOVA on the proportion of fixations from correct trials, with age (child, adult) as a between factor, and generalization distance (near, distant), time slice (beginning, middle, end), and AOI (Entity, Operator, Target Entity, Relational match, Thematically related, Taxonomically related) as within factors.

First, we considered results indicating participants' strategies' temporal dynamics (i.e., effects common to both age groups). Results reveal an effect of AOI, $F(5,455) = 173.90$, $p < .001$, $\eta^2_p = .66$, and an interaction between time slice and AOI, $F(10,910) = 153.36$, $p < .001$,

$\eta^2_p = .48$. This interaction is the most important for investigating strategies' temporal dynamics because it reveals the evolution of fixations patterns to different AOI over time. Post-hoc contrasts with a Bonferroni correction were run on this interaction. The interaction reveals that participants' attention is initially directed to learning items, followed by available options. Indeed, learning items are looked at more than all other items at the beginning of trials ($p < .001$); they are looked at more than available options in the middle of trials ($p < .001$) but as much as the Target Entity; and they are looked at less than the Target Entity and available options at the end of trials ($p < .001$). Available options are looked at equivalently except at the end of trials, where the relational match is looked at more than all other items ($p < .001$). This first result indicates that children focus on learning items at first and must extract information from these items, likely by comparing learning pairs. At this stage, we do not have any information about the type of strategy used. To determine this, we need to examine participants' eye movement patterns between items over time (i.e., described by switches).

Second, we considered results indicating the effect of generalization distance on strategy patterns (i.e., results including the generalization distance factor). Results revealed an interaction between generalization distance and AOI type $F(5,455) = 83.17, p < .001, \eta^2_p = .48$, and an interaction between generalization distance, time slice, and AOI type $F(10,910) = 105.30, p < .001, \eta^2_p = .54$ (Figure 2) that is the interaction that gives the most information about the effect of generalization distance on fixation times over time. We ran post hoc Bonferroni contrasts on this three-way interaction. The contrasts revealed that differences between patterns only appear at the end of trials. At the end of trials, in distant generalization cases, participants look more to learning entities and the relational match than they do in near generalization cases ($p < .001$). In near generalization cases, participants look more to the relational matches and the distractors than in distant generalization cases ($p < .001$).

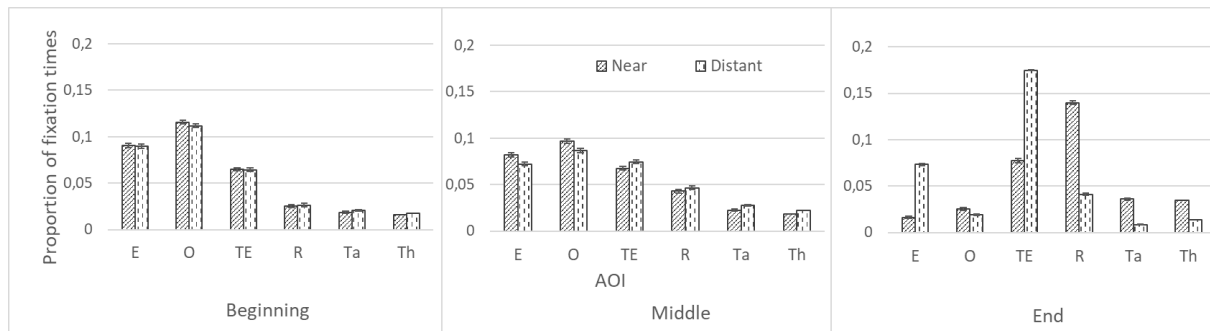


Figure 2: Proportion of fixation times as a function of generalization distance (near, distant), time slice (beginning, middle, end), and AOI (E, O, TE, R Ta, Th). Error bars are SEM.

Third, we considered results indicating strategy differences as a function of age (i.e., results including the factor of age). Results revealed an interaction effect between age and AOI, $F(5,455) = 12.47, p < .001, \eta^2_p = .12$, and a three-way interaction between age, time slice, and AOI, $F(10,910) = 17.12, p < .001, \eta^2_p = .05$ (see Figure 3). Post-hoc contrasts with a Bonferroni correction were run and revealed that differences between children’s and adults’ looking times appear in the beginning and at the end of trials, with the distribution of looking times being identical in the middle of trials. In the beginning of trials, adults look more to learning items than children ($p < .001$) and less to the TargetEntity than children ($p < .001$). Adults also look less to the distractors (i.e., Ta and Th) than children ($p < .001$). At the end of trials, adults look more to the relational match and less to the distractors than children ($p < .001$), revealing that they hesitate less between available options than children do.

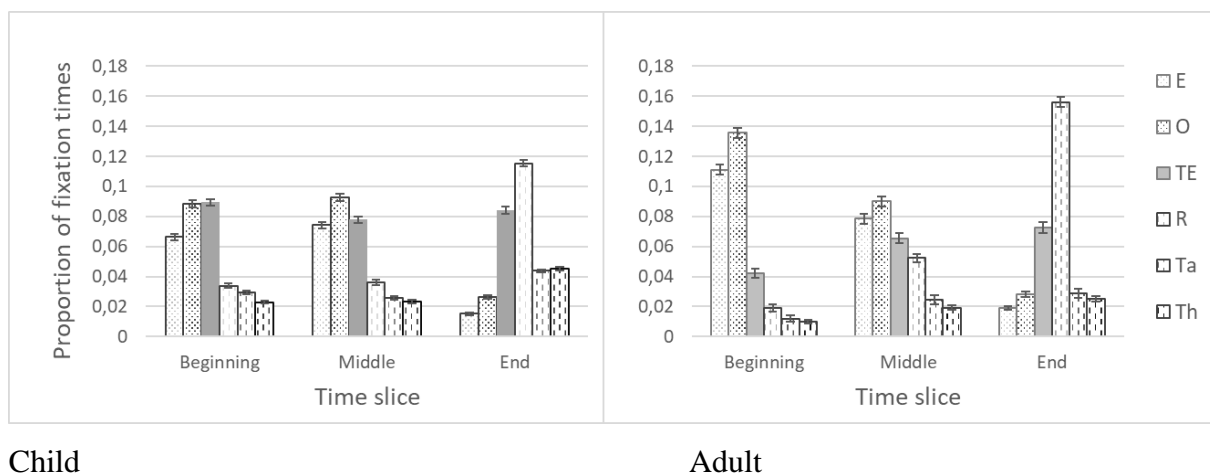


Figure 3. Proportion of fixations from correct trials as a function of age (child, adult), time slice (beginning, middle, end), and AOI (E, O, TE, R, Th, Ta). Error bars are SEM.

Results also revealed an interaction between generalization distance, age, time slice, and AOI type $F(10,910) = 3.65, p < .001, \eta^2_p = .04$. This interaction sums up previously

described effects, with a smaller effect size, in a more complex interaction. For these reasons, we chose to describe the conceptual distance and the age effects separately for clarity.

Switches

Switches were defined as movements from one AOI to another, provided children stayed at least 60 ms on each AOI. Since there were eight stimuli, which could give a priori 56 different switches, we aggregated them into eight different switch types that were the most meaningful for our analysis. Switches between learning items were aggregated into EOEO switches that are all switches between items of the same learning pairs (i.e., between entities E and operators O of a pair) or EEOO switches that are all switches between both of the trial's entities or both of the trial's operators. Any switches between a pair's entity and the other pair's operator were not considered. Switches including items from the trial's generalization domain (i.e., TE, R, Ta, Th) were aggregated into switches including only generalization domain items or between learning and generalization domain items. Switches in the generalization domain were TE-R switches between the target entity (i.e., TE) and the relational match (i.e., R), TE-Dis switches between TE and either of the distractors (i.e., Ta or Th), and Ch switches between the available choices (i.e., R, Ta, Th). Switches between learning and generalization domain items were E-TE switches between either of the learning entities and the target entity, O-R switches between either of the learning operators and the relational match and O-Dis switches between either of the learning operators and either of the distractors.

We analyzed the log transformation of the number of switches to fit ANOVA normality assumptions. We ran a three-way repeated measure ANOVA on data from correct trials, with age (child, adult) as a between factor and time slice (beginning, middle, end) and switch type (EOEO, TE-R, TE-Dis, Ch, EEOO, E-TE, O-R, O-Dis) as within factors.

First, we considered results indicating participants' strategies' temporal dynamics (i.e., results common to both age groups). Results revealed the main effects of the time slice, $F(2,210) = 14.46, p < .001, \eta^2_p = .12$, and switch type, $F(7,735) = 245.70, p < .001, \eta^2_p = .70$. Results also revealed an interaction between time slice and switch type, $F(14,1470) = 115.38, p < .001, \eta^2_p = .52$. This interaction is the most pertinent for investigating strategies' temporal dynamics because it reveals which switches were made and when during trials. Post hoc contrasts with a Bonferroni correction were run on this interaction and revealed that participants explored learning pairs before the TargetEntity and available options and that all

the most used switches are projection-type switches. Indeed, at the beginning of trials, participants made EOEO switches more than all other switches ($p < .001$). In the middle of trials, participants made EOEO and TE-R switches more than others ($p < .001$). At the end of trials, participants made TE-R, TE-Dis, and Ch switches more than all alignment type switches (i.e., EEOO, E-TE, O-R, O-Dis) and EOEO switches ($p < .001$).

Second, we considered results indicating the effect of generalization distance on strategy patterns (i.e., results including the generalization distance factor). Results revealed a main effect of generalization distance $F(1,105) = 5.43, p < .05, \eta_p^2 = .05$ showing that participants made more switches in distant generalization cases compared to near generalization cases ($M_{Near} = 0.29; M_{Distant} = 0.31$). Results also revealed an interaction between generalization distance, time slice, and switch type $F(14,1470) = 1.75, p < .05, \eta_p^2 = .02$. We ran post hoc Bonferroni contrasts. The contrasts, however, did not reveal any significant differences between strategies in near and distant cases.

Third, we considered results indicating strategy differences as a function of age (i.e., results including the factor of age). Results revealed a main effect of age, $F(1,105) = 37.5, p < .001, \eta_p^2 = .26$, where children made fewer switches than adults ($M_{child} = 0.27, M_{adult} = 0.38$); interaction effects between age and time slice, $F(2,210) = 12.55, p < .001, \eta_p^2 = .11$ and between age and switch type, $F(7,735) = 18.97, p < .001, \eta_p^2 = .15$; and a three-way interaction effect between age, time slice, and switch type and between age and time slice, $F(14,1470) = 18.20, p < .001, \eta_p^2 = .15$ (Figure 4). Post hoc contrasts with a Bonferroni correction were run on this last interaction, and we only considered differences between children and adults in the same time slice for the same switches. The interaction reveals that adults make more switches than children in their search. Indeed, if we consider projection type switches only to start with, contrasts reveal more EOEO switches for adults in the beginning ($p < .001$), more of all projection switches for adults in the middle of trials ($p < .001$), and more EOEO and TE-R switches for adults in the end of trials ($p < .001$). Interestingly, adults also make more alignment switches (i.e., EEOO, E-TE, O-R, O-Dis) than children from the middle of trials. These switches may be interpreted as used by adults to verify their choice. In the middle of trials, adults made more alignment switches than children ($p < .001$), and at the end of trials, they made more E-TE O-R and O-Dis switches than children ($p < .001$).

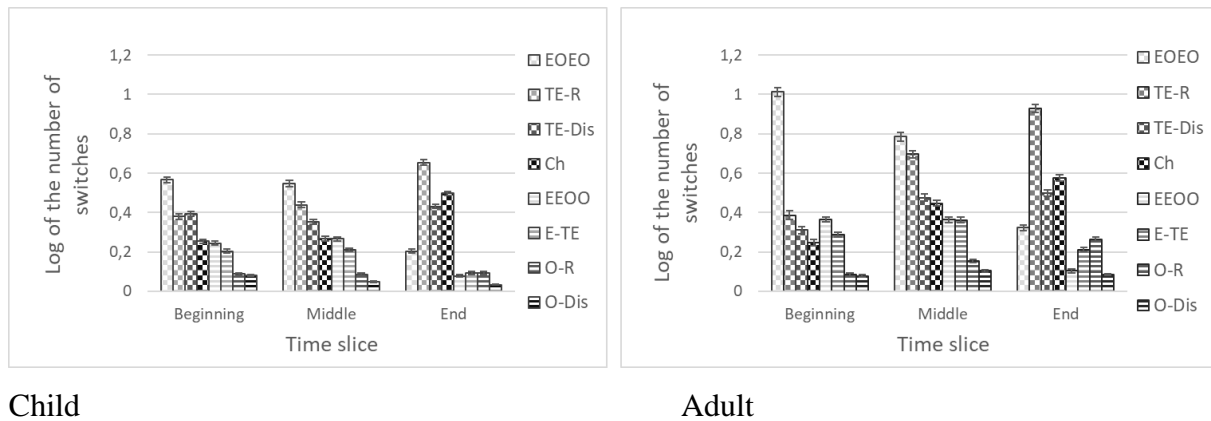


Figure 4. Log of the number of switches from correct trials as a function of age (child, adult), time slice (beginning, middle, end), and switch type (EOEO, TE-R, TE-Dis, Ch, EEOO, E-TE, O-R, O-Dis). Error bars are SEM.

Note: Projection-type switches are checked, and alignment-type switches are lined.

Results also revealed an interaction between generalization distance, age, time slice, and AOI type $F(14,1470) = 1.79, p < .05, \eta^2_p = .02$. This interaction sums up previously described effects, with a smaller effect size, in a more complex interaction. For these reasons and to ensure clarity, we chose to describe the conceptual distance and the age effects separately.

First gazes

The first analyses reveal that children look mostly to learning items (i.e., E and O) and the Target entity (i.e., TE) at the beginning of trials. However, the analysis of the number of switches reveals that they start looking from the target entity to the available option even at the beginning of trials. Hence, early differences in looking strategies, namely between learning items, may be obscured, as the first time slice has too broad a focus to identify early sequences and potential differences in strategies. To obtain finer-grained data, we considered children's first five gazes and analyzed what items they looked at first. Our research hypothesizes that participants' strategies will have an alignment-first, a projection-first, or a goal-directed pattern. Previous research indicates that a projection-first pattern involving eye movements between the TargetEntity and available options is likely. A growing body of research indicates that time learning before moving on to the generalization search is a crucial task step. However, this implies sufficient task control to focus on the learning items before the immediate goal. This more efficient strategy, where goal resolution is delayed, likely

involves better executive skills. We, therefore, predict that children’s ability to delay the tasks’ goal (i.e., search for a generalization solution) will be less efficient than adults’ and that this will be reflected in their early gaze orientations.

For each of the first five gazes, we calculated the average proportion of times they looked to a learning item, the target entity, or one of the available choices.

We ran a three-way repeated measure ANOVA on the proportion of gazes from the first five gazes from correct trials with age (child, adult) as a between factor and gaze location (learning, target entity, available choices) and gaze number (1, 2, 3, 4, 5) as within factors.

First, we considered results indicating participants' strategies' temporal dynamics (i.e., results common to both age groups). Results revealed a main effect of gaze location $F(2,186) = 198.26, p < .001, \eta^2_p = .68$ that confirms that participants look more to learning items than the target entity and available choices across the first five gazes.

Second, we considered results indicating strategy differences as a function of age (i.e., results including the factor of age). Results revealed an interaction between age and gaze location $F(2,186) = 79.17, p < .001, \eta^2_p = .46$, and a three-way interaction between age, gaze location, and gaze number $F(8,744) = 12.34, p < .001, \eta^2_p = .12$ (Figure 5). Post hoc contrasts with a Bonferroni correction were run on this last interaction. These contrasts reveal that, for all five first gazes, children looked less at the learning items and more at the available choices than adults ($p < .001$). They also looked more at the target entity than adults in the first and third gaze ($p < .001$).

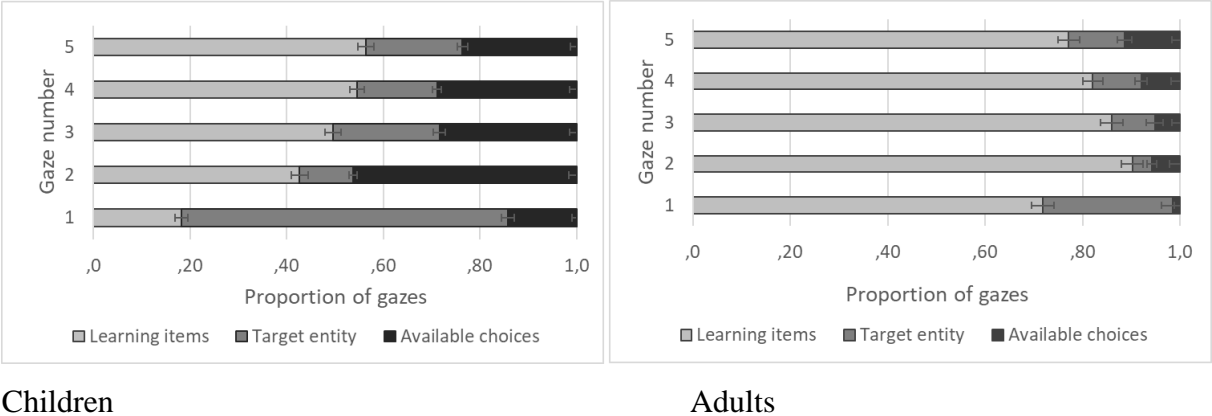


Figure 5. Proportion of fixations of first five trials from correct trials as a function of age (child, adult), gaze location (learning, target entity, available choices), and gaze number (1, 2, 3, 4, 5). Error bars are SEM.

*Eye-tracking data: Strategy variations*Strategy variations as a function of trial outcome

One of our questions was how strategy implementation may determine success. One way to study this is to compare gaze profiles from trials that led to taxonomic answers (i.e., “correct” answers) to gaze profiles from trials that led to other answers (i.e., “errors”), using a *trial accuracy* factor as a between factor. This should reveal how children come up with a solution, how correct and error profiles differ, and if the differences between these gaze profiles, if any exist, are mainly at the beginning or the end of trials. Thus, we are mainly interested in interactions, including the factor trial accuracy.

The following analyses are done with children’s data only because the proportion of error trials is too low in adults for such an analysis.

Proportion of fixation times

We ran a three-way repeated measure ANOVA on the proportion of fixations from all children trials, with trial accuracy (correct, error), time slice (beginning, middle, end), and AOI (E, O, TE, R, Th, Ta) as within factors. Here, only the effects including trial accuracy are those of interest. Results revealed interaction effects between trial accuracy and AOI $F(5,295) = 22.39$, $p < .001$, $\eta^2_p = .28$, and between trial accuracy and time slice $F(2,118) = 6.63$, $p < .001$, $\eta^2_p = .10$. The analysis also revealed a three-way interaction between trial accuracy, time slice and AOI $F(10,590) = 21.61$, $p < .001$, $\eta^2_p = .27$ (see Figure 6).

The three-way interaction with trial accuracy, time slice, and AOI is the effect of most interest here. This is the one to be represented (see Figure 6) and discussed because it gives the best insight into search strategies’ temporal dynamics as a function of trial accuracy. We ran post hoc contrasts for this interaction, with a Bonferroni corrected significance level set at $p = 0.05/18 = 0.003$, because only the eighteen contrasts between correct and error profiles for the same AOI in the same slices are of interest here. These contrasts reveal that differences between correct and error trials appear at the end of trials, where R is looked at more in correct trials than in error trials and where Th and Ta are looked at more in error trials than in correct trials.

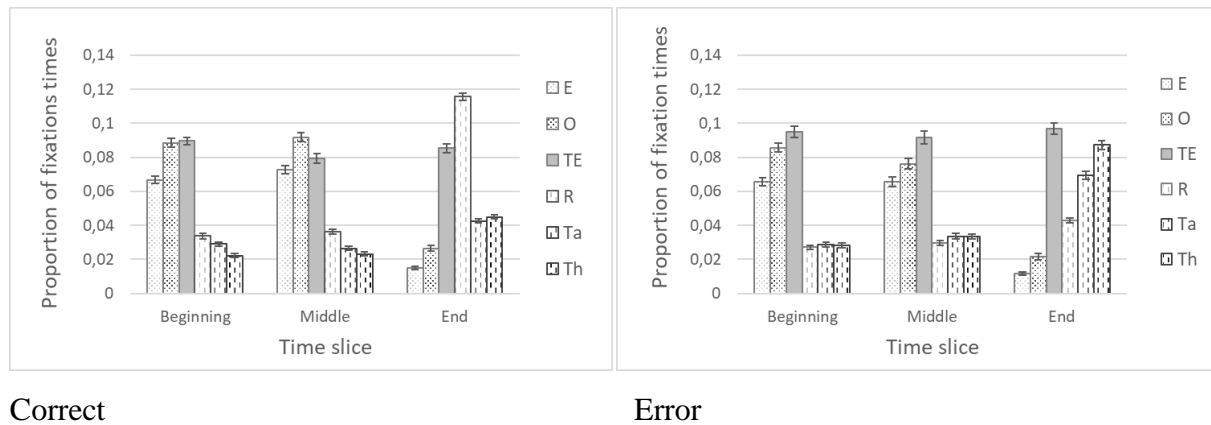


Figure 6. Proportion of fixations from all trials as a function of trial accuracy (correct, error), time slice (beginning, middle, end), and AOI (E, O, TE, R, Ta, Th). Error bars are SEM.

Number of switches

We ran a three-way repeated measure ANOVA on the number of switches' logs from all children trials, with trial accuracy (correct, error), time slice (beginning, middle, end), and switch type (EOEO, TE-R, TE-Dis, Ch, EEEO, E-TE, O-R, O-Dis) as within factors. Here, the effects including trial accuracy are those of interest. Results revealed a main effect of accuracy $F(1,59) = 16.95, p < .001, \eta^2_p = .22$, interaction effects between accuracy and time slice $F(2,118) = 3.70, p < .05, \eta^2_p = .06$, and between trial accuracy and switch type $F(7,413) = 17.63, p < .001, \eta^2_p = .23$. The analysis also revealed a three-way interaction between trial accuracy, time slice and $F(14,826) = 11.85, p < .001, \eta^2_p = .17$ (see Figure 7).

The three-way interaction with trial accuracy, time slice, and switch type is the effect of most interest here and the one to be represented (see Figure 7) and discussed because it gives the best insight into search strategies' temporal dynamics as a function of trial accuracy.

We ran post hoc contrasts for this interaction, with a Bonferroni corrected significance level set at $p = 0.05/24 = 0.002$, because only the twenty-four contrasts between correct and error profiles for the same switches in the same slices are of interest here. These contrasts reveal that, in the beginning, middle, and end of trials, children made more EOEO switches in correct trials rather than trials that led to errors ($p < .002$). Children also made more TE-R switches in correct rather than error trials in the middle and end of trials ($p < .001$). At the end of trials, children also made more Ch switches in correct rather than error trials ($p < .001$).

Concerning switches considered alignment switches, children made more EEOO switches in correct trials rather than error trials ($p < .001$) in the beginning and middle of trials. And more TE-E switches in the middle and end of trials ($p < .001$).

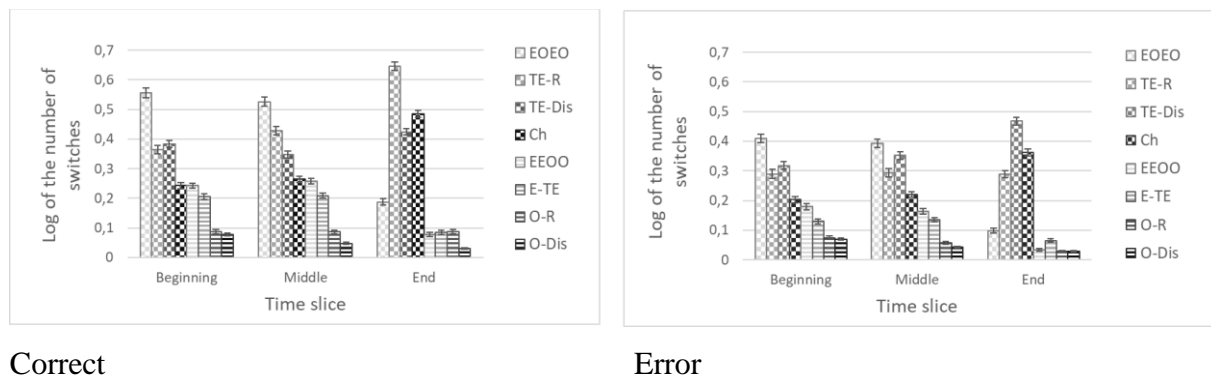


Figure 7. Log of the number of switches from all trials as a function of trial accuracy (correct, error), time slice (beginning, middle, end), and switch type (EOEO, TE-R, TE-Dis, Ch, EEOO, E-TE, O-R, O-Dis). Error bars are SEM.

Note: Projection-type switches are checked, and alignment-type switches are lined.

First gazes

We ran a two-way repeated measure ANOVA on the proportion of gazes from the first five gazes from all children's trials with gaze location (learning, target entity, available choices) and gaze number (1, 2, 3, 4, 5) as within factors. Here, the effects including trial accuracy are those of interest. However, the analysis did not reveal any effects of accuracy on children's first gazes. For example, the interaction between accuracy, gaze location, and gaze number would have revealed differences in children's early gaze orientations as a function of accuracy for all five early gazes. However, it was non-significant $F(8,464) = 1.02, p = 0.42, \eta^2_p = .02$. Whether their final answer was a correct answer or an error, children always looked mainly to the target entity in their first gaze, and their attention was then divided between learning and generalization items.

Discussion:

Word generalization with comparison settings at learning is described as mapping tasks between items, mental representations, and prior knowledge. However, little is known about how the mapping occurs and how participants achieve correct generalization from the information they process in comparison and generalization tasks. Our primary goal in the present study was to describe the temporal dynamics of relational noun generalization

strategies in comparison settings. We proceeded in three steps. First, by (1) describing the strategies' temporal dynamics and interpreting them in terms of existing strategies. This is followed by two points that aim at getting an insight into what may be determinant in the strategies for generalization success by (2) assessing how strategies may develop and (3) investigating differences between strategies that lead to different outcomes. Overall, our results reveal that both adults' and children's switches and fixation times reflected a projection-first strategy in which they considered learning pairs as wholes before matching available options to the given Target Entity. The main difference observed between adults and children was that at the onset of trials, children turn to generalization items more than adults, who focus solely on learning items before turning to generalization items. We will discuss the implications of this difference for generalization success.

Interpreting eye movements as strategies

Our first and foremost aim was to describe relational noun strategies' temporal dynamics and interpret them in light of given strategies, the alignment, projection, and goal-directed strategies. We analyzed three different eye-tracking indexes: proportion of fixation times to AOI, switches between AOI, and first gaze location. For all indexes, both children's and adults' data are consistent with a projection strategy hypothesis that predicts early fixations to learning items and considers learning pairs as wholes to extract a representation of the targeted relation, then turns to generalization items to project the extracted relation in the generalization domain, and finds a solution by sequentially aligning the given target entity to the available option to choose the pair that best fits the mental representation. An alignment strategy predicted switches between equivalent items in the learning and the generalization domains. The data reveals that participants make few of these types of switches. We interpret that the alignment type switches recorded are noise obtained from the search rather than switches with meaning in terms of participants' search. The projection-type strategy is consistent with the strategies hypothesized (Hummel & Holyoak, 1997) or revealed (Starr et al., 2018; Vendetti et al., 2017) in analogical reasoning. It is interesting to see that the projection strategy describes both analogical reasoning and relation noun generalization tasks.

Further, than simply identifying the strategy type used by participants in relational noun generalization, we were particularly interested in the strategy's temporal dynamics. This question is crucial to understand which steps are needed to reach generalization, which cognitive process may be involved, and, as discussed below, which steps are determinants for generalization success. When considered for all participants (i.e., children and adults), the

data reveals a remarkably clear-cut sequence of three steps. First is a learning step: fixations at the beginning of trials are directed toward learning items, and switches mainly occur between learning items. Second, a solution step: in the middle of trials, fixations are directed toward learning items and the target entity, around which the search for the solution must be organized, and switches between the target entity and the solution appear. Third and finally, a verification step: at the end of trials, participants focus on all generalization items (i.e., target entity and all available options) instead of learning items. The dominant switches solely occur in the generalization domain, involving all items.

This clear-cut sequence supports the hypothesis behind the projection strategy. Indeed, the hypothesis is that a representation of the targeted relation is extracted from the learning setting and projected in the generalization domain. Our data reveals a clearly identifiable learning phase coherent with the need to consider learning items deeply to extract a relation. A solution is then rapidly found in the middle of trials, which supports that the solution is found as a match for the representation rather than progressively built by combining task elements. What our data adds is the verification phase at the end of trials. This verification phase stays consistent with the projection strategy's structure because all verification switches are solely in the generalization domain. Verification could be interpreted as a verification between the generalization items and the mental representation rather than between the learning examples and the solution, which is what would best fit the alignment strategy's structure.

The above description of the data as a whole, which merges children's and adults' data, is enlightening when looking for relational noun generalization strategies because it reveals both strategy types and a clear description of the strategy's temporal dynamics. However, it does not give the whole scope of the data, which only appears when children's and adults' strategies are confronted. In the following section, we shall discuss the differences between children's and adults' strategies and how these differences indicate which strategy steps may be determinant for generalization success.

Strategy development and the differential importance of strategy steps

We compared children's and adults' strategies, and the main difference that appeared was at the beginning of trials. For all three eye-tracking indexes, adults focused more on learning items at the beginning of trials than children. More precisely, attention to the trials' target entity interfered with children's fixations, switches, and first glances at learning items

at the beginning of trials. This is consistent with the goal-directed strategy (Thibaut & French, 2016), which describes that children's search is diverted towards that task's goal, thus impeding their learning phase. From an executive point of view, children's behavior may be caused by a still-developing inhibition ability, whereas adults have fully developed executive functions (Diamond, 2013) that allow them to avoid this flaw.

One can mention that the differences between children and adults appear to be more of efficiency rather than structure. This is to say that, even at a young age (i.e., here, five years old), children's strategies are those of fully developed adults. Therefore, what children are tackling during their own development and relational noun generalization development is not acquiring the right strategy but learning to control their strategy most efficiently. In other words, we consider that the performance differences revealed by our behavioral data are due to adults' better control of their strategy compared to children's rather than a different strategy in itself.

Finally, we stress the importance of the beginning of trials and children's ability to focus on the source of information before searching for a solution. At this stage, one could consider that the key to successful generalization is in the beginning. This is correct but deserves to be nuanced, as shown by the last analysis comparing strategies from trials with different outcomes.

Strategy implementation and consequences for generalization outcome

Our last purpose was to understand how strategy implementation affected generalization success. The comparison between adults' and children's data partially answers this question by revealing the importance of the learning phase at the beginning of trials. The comparison of strategies that led to correct answers or errors was another way to address this problem. Data reveals that the differences in children's strategies from correct vs. error trials are at the end of trials, that is, in the decision-making to find a solution. At first glance, this contradicts the above developmental analysis, which points to the beginning of trials as the critical step for generalization success. We will now discuss how these two suggestions can be considered complementary.

Since participants use a projection-type strategy to generalize novel relation nouns in the present task, they must be able to match their representation of the relation to a solution in the generalization domain. We suggest that children who built a poor representation struggle to match a solution to their representation because their learning phase was impoverished by

attention to the target entity. Therefore, children may hesitate when deciding at the end of trials.

This result argues for the importance of the mental representation built during learning. It appears that it might play a decisive part in generalization success. If this were to be the case, a question could be raised regarding the effect of semantic distance. We did not reveal an effect of semantic distance on generalization search strategies. This is most likely due to the ease of the task for all participants. However, in different conditions, would semantic distance affect the strategies or solely the representation (Holyoak, 2012) that eye-tracking data cannot reveal?

Limits of the present study

Relational noun generalization has previously been studied with concluding data in children as young as three (Gentner et al., 2011b). Considering that children's percentage of relational choices was relatively high and may have impeded revealing an effect of semantic distance between items, it is necessary to study younger children's relational noun generalization strategies. The present equipment makes this too challenging. Indeed, with a Tobii 120 device, 4-year-old children move too much for productive data collection. However, equipment such as eye-tracking glasses may be considered.

One may also question the influence of the task's layout on eye movements, thus on the conclusion about generalization strategies. Points of control and adaptations were brought to tackle this critical question. No layout effect was found among the four counterbalanced layouts across participants. Layouts were designed to counter the effects of reading directions. Our display avoided displaying adjoining items in learning pairs as it may be found in the literature, maximizing the distance similarity between the same learning pair items and different learning pair items. However, the influence of the display on eye movements is an ongoing concern. Now that the first exploratory study into relational noun generalization strategies has been made and that it has been shown that children and adult strategies are of the same type, it would be of interest to suggest tasks with more complicated layouts and older children. It would allow us to investigate if a projection strategy is still used. Following Bethell-Foxes et al. (1984b), one may discover that as the task's design evolves, the constraints it entails on participants' search will translate as switches between different types of strategies.

Concluding remarks

The present study gives a first description of relational noun generalization strategies and reveals that in both adults and children, these strategies are consistent with a projection-type strategy. The data also reveals that generalization strategies are acquired at a young age and become progressively more efficient with increasing age. Questions regarding the importance of the representation built during the mapping process and of the task's constraints arise from the present work and open interesting possibilities for future investigation into children's conceptual development.

References

- Baldwin, D. A. (1995). Understanding the link between joint attention and language. *Joint attention: Its origins and role in development*, 131, 158.
- Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984). Adaptive reasoning : Componential and eye movement analysis of geometric analogy performance. *Intelligence*, 8(3), 205-238.
[https://doi.org/10.1016/0160-2896\(84\)90009-6](https://doi.org/10.1016/0160-2896(84)90009-6)
- Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
- Carey, S., & Bartlett, E. (1978). *Acquiring a single new word*.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135-168.
<https://doi.org/10.1146/annurev-psych-113011-143750>
- Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure Mapping and Relational Language Support Children's Learning of Relational Categories : Structure Mapping and Relational Language. *Child Development*, 82(4), 1173-1188.
<https://doi.org/10.1111/j.1467-8624.2011.01599.x>
- Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist*, 52, 45-56.
- Gentner, D., & Namy, L. L. (2004). The Role of Comparison in Children's Early Word Learning. In *Weaving a lexicon* (p. 533-568). MIT Press.
- Gentner, D., Rattermann, M. J., & Forbus, K. D. (1993). The roles of similarity in transfer : Separating retrievability from inferential soundness. *Cognitive psychology*, 25(4), 524-575.

- Gentner, D., & Toupin, C. (1986). Systematicity and Surface Similarity in the Development of Analogy. *Cognitive Science*, *10*(3), 277-300.
https://doi.org/10.1207/s15516709cog1003_2
- Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive psychology*, *15*(1), 1-38.
- Holyoak, K. J. (2012). Analogy and Relational Reasoning. In K. J. Holyoak & R. G. Morrison (Éds.), *The Oxford Handbook of Thinking and Reasoning* (1^{re} éd., p. 234-259). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199734689.013.0013>
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure : A theory of analogical access and mapping. *Psychological review*, *104*(3), 427.
- Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological review*, *110*(2), 220.
- Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, *8*(2), 113-139.
- Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.
- Kucker, S. C., Samuelson, L., Perry, L., Yoshida, H., Colunga, E., Lorenz, M., & Smith, L. (2019). Reproducibility and a Unifying Explanation : Lessons from the Shape Bias. *Infant behavior & development*, *54*, 156-165. <https://doi.org/10.1016/j.infbeh.2018.09.011>
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development*, *3*(3), 299-321.
- Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. *Trends in Cognitive Sciences*, *4*(1), 6-14. [https://doi.org/10.1016/S1364-6613\(99\)01418-7](https://doi.org/10.1016/S1364-6613(99)01418-7)
- Markman, A. B., & Gentner, D. (1993a). Splitting the Differences : A Structural Alignment View of Similarity. *Journal of Memory and Language*, *32*(4), 517-535.
<https://doi.org/10.1006/jmla.1993.1027>

- Markman, A. B., & Gentner, D. (1993b). Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, 25(4), 431-467.
<https://doi.org/10.1006/cogp.1993.1011>
- Markman, A. B., & Gentner, D. (1996). Commonalities and differences in similarity comparisons. *Memory & Cognition*, 24(2), 235-249. <https://doi.org/10.3758/BF03200884>
- Markman, E. M. (1989). *Categorization and naming in children : Problems of induction*. MIT Press.
- Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological review*, 100(2), 254.
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Rayner, K., Shen, D., Bai, X., & Yan, G. (2018). *Cognitive and cultural influences on eye movements*. CRC Press.
- Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning : Insights from scene analogy problems. *Journal of Experimental Child Psychology*, 94(3), 249-273. <https://doi.org/10.1016/j.jecp.2006.02.002>
- Smith, L. B. (1999). Children's noun learning : How general learning processes make specialized learning mechanisms. *The emergence of language*, 277-303.
- Soluch, P., & Tarnowski, A. (2013). Eye-tracking methods and measures. *Translation Studies and Eye-Tracking Analysis. Frankfurt am Main: Peter Lang*, 85-104.
- Starr, A., Vendetti, M. S., & Bunge, S. A. (2018). Eye movements provide insight into individual differences in children's analogical reasoning strategies. *Acta Psychologica*, 186, 18-26. <https://doi.org/10.1016/j.actpsy.2018.04.002>
- Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, 27(2), 195-232. [https://doi.org/10.1016/0022-0965\(79\)90044-4](https://doi.org/10.1016/0022-0965(79)90044-4)

- Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions : A developmental investigation with eye-tracking. *Cognitive Development*, 38, 10-26.
- Thibaut, J.-P., French, R., & Vezneva, M. (2010). The development of analogy making in children : Cognitive load and executive functions. *Journal of Experimental Child Psychology*, 106(1), 1-19. <https://doi.org/10.1016/j.jecp.2010.01.001>
- Thibaut, J.-P., & Witt, A. (2015). Young children's learning of relational categories : Multiple comparisons and their cognitive constraints. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00643>
- Thibaut, J.-P., & Witt, A. (2017). Generalizing novel names in comparison settings : Role of conceptual distance during learning and at test. *CogSci 2017 : 39th Annual Meeting of the Cognitive Science Society*, 3314-3319. <https://hal-univ-bourgogne.archives-ouvertes.fr/hal-01573610>
- Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, 10(2), 201-224. [https://doi.org/10.1016/0885-2014\(95\)90009-8](https://doi.org/10.1016/0885-2014(95)90009-8)
- Vendetti, M. S., Starr, A., Johnson, E. L., Modavi, K., & Bunge, S. A. (2017). Eye Movements Reveal Optimal Strategies for Analogical Reasoning. *Frontiers in Psychology*, 8. <https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00932>
- Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations : Toddlers' and preschoolers' word-learning biases. *Child development*, 61(5), 1461-1473.
- Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, 114(2), 245-272. <https://doi.org/10.1037/0033-295X.114.2.245>

Chapter 3 – Investigating cognitive mechanisms that underlie novel word generalization.

The following work's reflexional is of interest even though its results were not conclusive. It is not aimed for publication but is nonetheless presented in the following chapter.

Introduction

A large body of research has led to a good understanding of children's categorical choices in novel word generalization. This research has revealed how children's early generalizations are biased by the input's properties. Perceptual similarity, for example, can lead children to choose perceptually similar items instead of conceptually related items (Baldwin, 1989; Graham et al., 2010; Imai et al., 1994a; Landau et al., 1988b). Items previously experienced to have the same function or to be otherwise related can also lead children to extend words to items that are not conceptually related (Ware, 2017).

Helping children overcome these biases to achieve taxonomically based word generalization is crucial for correct word use and efficient communication. Research has consequently investigated which learning and generalization situations favor conceptually based word extensions over other extensions. Many task parameters have been shown to influence word generalization in children such as number of learning items (Augier & Thibaut, 2013; Gentner et al., 2011b; Price & Sandhofer, 2021), conceptual distance between learning items or between learning and generalization items (Thibaut et al., 2018; Thibaut & Witt, 2015), learning item's perceptual similarity (Namy et al., 2007b), the presence of a common item label (Namy & Gentner, 2002), and the item's availability at learning (Kovack-Lesh & Oakes, 2007; Oakes & Ribar, 2005) or at generalization (Stansbury et al., 2022).

However, this approach is solely category-based. It is informative of which categories are followed in which learning situations. However, it says very little about how children come to a generalization decision. The underpinning mechanisms of word generalization have at best been inferred from the variations noticed in children's generalization behavior (Alfieri et al., 2013a; Augier & Thibaut, 2013; Hammer et al., 2008; Markman & Gentner, 1993b), but they have often been left aside. It therefore seems necessary to shift the focus towards the cognitive processes that support novel word generalization.

Mechanisms underlying novel word generalization

Novel word generalization is a step in the word learning process. Learning new words to generalize them involves processing the external stimuli – the word-item co-occurrence noticed in the environment – and adapting the existing conceptual knowledge to the newly experienced information (i.e., integrating the word to the concept, adding newly noticed features to the items' representation, reweighting the representation's features, etc.).

Processing the word-item co-occurrence first involves multimodal processing of the words and the item's features. Then, learners must selectively focus on the most relevant item features (i.e., the category membership features) to be abstracted and added to the existing conceptual knowledge with the newly learned noun. Word generalization targets will thereafter be the items that fit the built representation.

In the present work, we will concentrate on the processes that support processing the environmental stimuli and finding the generalization target. These processes enable learners to efficiently direct their selective attention to the item's relevant features and search and choose the generalization target.

Searching for the inputs' relevant features most likely involves spatial intelligence. Indeed, spatial intelligence is the ability to accurately perceive the visual and spatial world and transform those perceptions into concepts. This ability seems perfectly adapted for early concept building that involves processing and memorizing perceptually grounded features but may quickly become limited for building the more abstract concepts that structure human conceptual knowledge.

Novel noun generalization involves concepts with abstract features, such as words and conceptual features. As relational noun generalization involves relational features, focusing on these less salient features, often unrelated to perceptually grounded features, may depend on the maturity of executive functions (Sloutsky, 2010). Switching away from the item's irrelevant dimensions (e.g., perceptually grounded features) to more relevant but less salient dimensions (e.g., conceptual features) may depend on cognitive flexibility. In the same way, switching from solely perceptually grounded representations to conceptual representations (Blaye & Bonthoux, 2001; Blaye & Jacques, 2009) and mapping unobservable dimensions to stored abstract features may involve mature cognitive flexibility and working memory.

An executive framework for novel word generalization

The executive function framework suggests that processing the generalization tasks' input involves searching the task's space to compare multiple items and choose a match while neglecting irrelevant dimensions, such as perceptual similarities, to focus on more relevant conceptual dimensions. A growing body of research suggests that these processes generate cognitive cost (Richland et al., 2006; Thibaut et al., 2010a) because of holding information in working memory (Diamond, 2013), switching between items, decision-making, and inhibition. Therefore, they suggest the preponderant role of executive functions both in

directing participants' attention during the task and in their generalization decision-making. Results interpreted in this executive functioning framework have corroborated its suggestion in analogical reasoning tasks (Thibaut et al., 2010), novel noun generalization (Lagarrigue & Thibaut, 2020; Thibaut & Witt, 2015, 2017); relational noun generalization (Gentner et al., 2011b); noun generalization with unknown stimuli (Augier & Thibaut, 2013; Lagarrigue, 2022).

Augier and Thibaut (2013) studied the conceptualization of unfamiliar objects in a comparison paradigm and manipulated the number of exemplars shown during the comparison phase. They tested 4- and 6-year-olds and compared a no-comparison condition, a two-item comparison condition, and a four-item comparison condition. Interestingly, all children benefited from the comparison conditions compared to the no-comparison conditions. However, only older children benefited from the four-item comparisons compared to the two-item comparisons. This suggests that cognitive control is necessary to succeed in the task, as suggested by contributions in numerous domains involving comparison and integration of multiple information (see Wiebe & Karbach, 2017).

The semantic distance between the compared items might contribute to increase the cognitive costs of comparison. For example, Green et al. (2010) have shown that analogies based on distant domains were more difficult than equivalent analogies connecting closer domains because distant analogies involved more creativity, which was related to the central role of the prefrontal cortex in cognitive control. In children, Thibaut, French, and Vezneva (2010a) have shown that semantic analogies based on weakly associated relations are more difficult than those based on strongly associated relations. The authors interpreted this result in terms of the necessity to inhibit strongly associated but irrelevant items in the context at hand or in terms of the necessity to generate new candidate relations, which requires cognitive flexibility in the case of distant semantic domains.

The drawback of these studies is that they only consider participants' generalization choices as a function of the learning settings. Thus, they do not allow a direct link between individual executive functioning development and the generalization of analogical reasoning answers. Simms et al. (2018) adopted a different approach by investigating how working memory, inhibitory control, and cognitive flexibility measures predicted 5- to 11-year-old children's performance on an analogical mapping task. Cognitive function measures were taken using tasks from the NIH Toolbox Cognition Battery (2013). Their results revealed that individual differences in children's working memory best predicted their analogy

performances. Even after controlling for age, the relationship remained significant, suggesting a strong interrelationship between analogical reasoning development and working memory development.

Lagarrigue & Thibaut (2020) adopted the same approach regarding novel noun generalization with both unknown and familiar items. They tested 4- to 5-year-old children in a comparison and generalization task in which the task demands were manipulated to vary the task's cognitive cost (i.e., semantic distance between items for familiar objects, perceptual distinctiveness for unknown items). They also measured executive function levels with tasks adapted from the NIH Toolbox Cognition Battery (2013). They controlled for the role of conceptual knowledge by assessing children's level of vocabulary as an indicator of their general knowledge of the world. Their results revealed positive correlations between cognitive flexibility and novel noun generalization performances for unknown items and correlations between both cognitive flexibility and inhibition for familiar items. Interestingly, cognitive flexibility was strongly correlated to the other executive function measures (i.e., working memory and inhibition for unknown items, working memory for familiar items). Moreover, in linear mixed models, cognitive flexibility was revealed as a predictor for generalization performances in both types of generalization tasks.

More surprisingly, cognitive flexibility was also correlated to vocabulary measures in familiar items' noun generalization but also in unknown items' noun generalization. This implies that conceptual knowledge is likely used in any novel knowledge acquisition task. This finding is supported by the knowledge view of conceptual development that suggests that any type of concept building is consistent and intertwined with all existing conceptual knowledge structures (Keil, 1992; Murphy & Medin, 1985). The dominant idea behind this suggestion is that people pay attention to the features that their prior knowledge says are important.

The present studies

In the following section, we will describe our exploration of the processes that support novel relational noun and noun generalization. We adopted the same approach as Simms et al. (2018) and Lagarrigue et al. (2020), measuring cognitive processing levels (i.e., spatial intelligence, executive functions, and existing knowledge) for participants who took part in the novel noun or novel relational noun generalization tasks described in chapters 2 and 3 of this section.

We assessed inhibition, working memory, cognitive flexibility, and vocabulary for participants tested for novel relational noun generalization and inhibition, working memory, spatial intelligence, and vocabulary for participants tested for novel noun generalization.

Our main suggestion is that novel word comparison and generalization tasks are supported by multiple processes crossing information processing mechanisms (e.g., executive functions, spatial intelligence) and existing conceptual knowledge.

We therefore suggest that vocabulary, as an indicator of children's level of knowledge of the world, will be predictive of children's generalization performances (Keil, 1992; Murphy & Medin, 1985).

We suggest that executive functions are involved in the processes of comparison and generalization decision-making to inhibit automatic answers based on salient but irrelevant dimensions, to switch between items flexibly or to re-represent the stimuli along dimensions that did not come to immediately, and to efficiently manipulate the stimuli's input and map it to existing knowledge in working memory. We expect that poorer executive functioning levels will lead to more generalization errors.

Finally, we also tested spatial intelligence for children who participated in the novel noun generalization task. The structural alignment hypothesis (Markman & Gentner, 1993b) suggests that aligning perceptual similarities in learning item sets enhances the search for deeper conceptual similarities. In this case, children's ability to search the perceptual space efficiently may be an a priori requisite for successful comparison leading to the detection of conceptual similarities between learning items.

Method

Participants

Seventy-six children participating in the relational noun generalization task were tested for vocabulary, working memory, flexibility, and inhibition. Children were 5 years old (mean age: 66 months, range: 56 to 74 months).

One hundred and sixty-two children who had participated in the noun generalization task were tested for visual-spatial reasoning, vocabulary, working memory, and inhibition. Children were from three different age groups: forty-three children were five years old (mean age: 63 months, range: 58 to 69 months), sixty-four were six years old (mean age: 80 months,

range: 74 to 87 months), and one hundred were eight years old (mean age: 100 months, range: 91 to 116 months).

All the children were native French speakers. Informed consent was obtained from their school and their parents. The procedure was in accordance with the Declaration of Helsinki (1964). It was realized in the context of a research convention (convention n°: 0482-2021) between the laboratory, the university, and the academic inspectorate (“Inspection Académique de Côte d’Or”).

Cognitive assessments

In the relational noun generalization task vocabulary, three executive functions (flexibility, working memory, inhibition) were assessed. In the noun generalization task, vocabulary, visual-spatial reasoning, and two executive functions (working memory, inhibition) were assessed. The tests used are described below.

Vocabulary

For the vocabulary test, we used the EVIP, which is a French adaptation (Canadian norms) of the PPVT (Peabody Picture Vocabulary Test, (*Peabody Picture Vocabulary Test--Third Edition*, Dunn & Dunn, 1997). In this test, children had to choose an image from a selection of four pictures on a card (See Figure 1). The task was progressively more difficult and stopped when children gave fewer than six correct answers in eight trials.

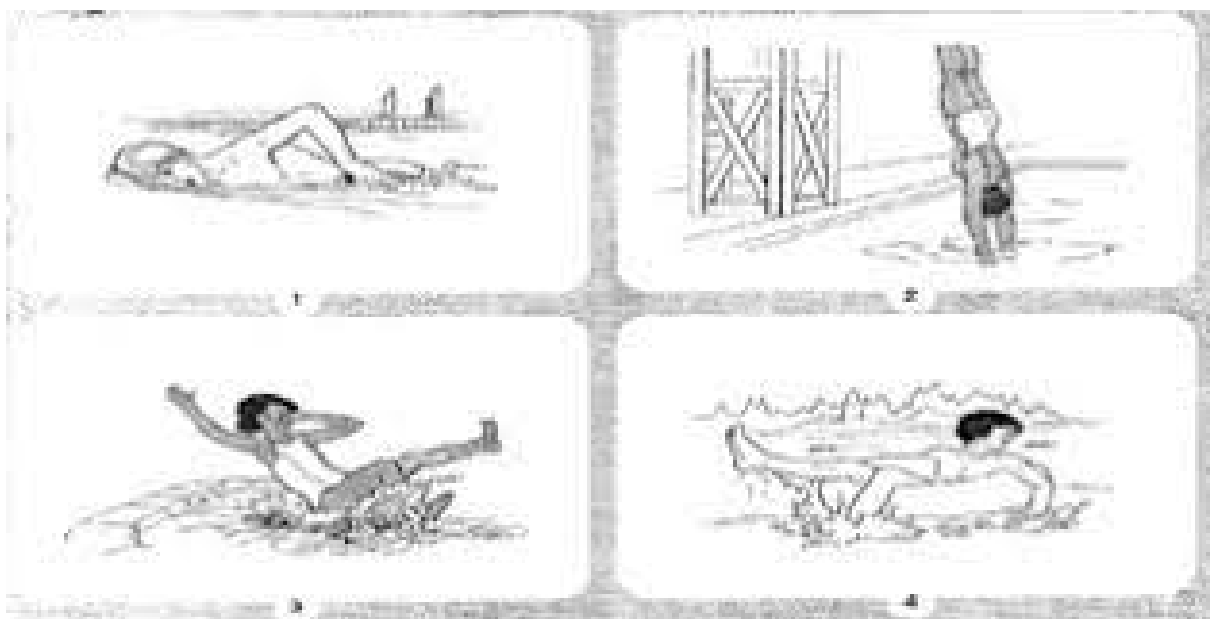
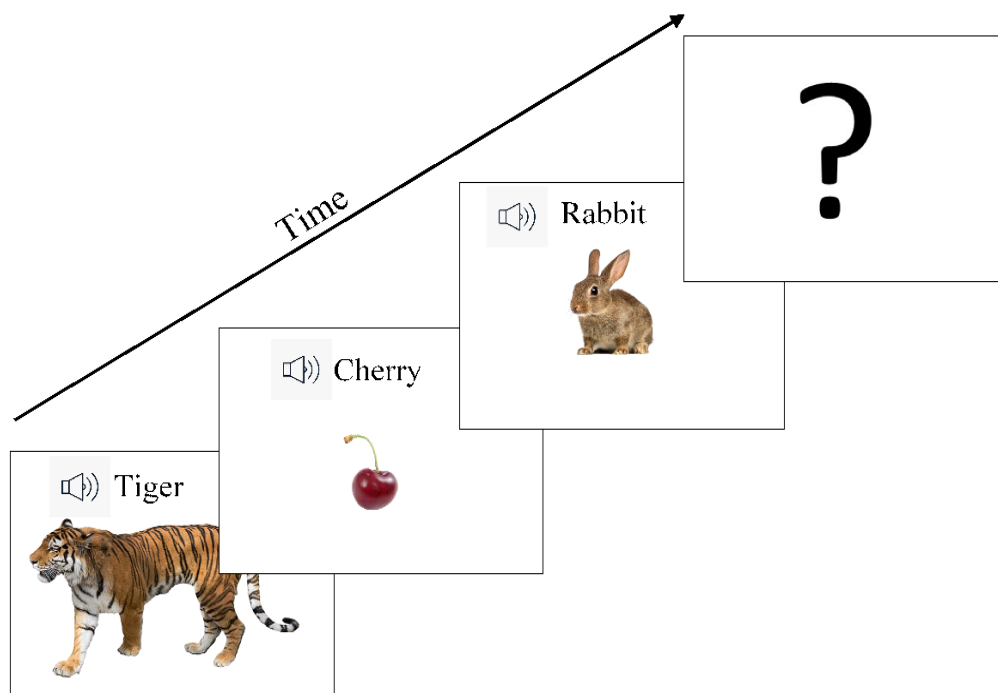


Figure 1: EVIP card example shown with the word “to dive.”

Working memory

We adapted the List Sorting Working Memory task from the National Institutes of Health Toolbox Battery (NIH Toolbox CB). We followed the same protocol, except that we implemented the tasks on OpenSesame, and the instructions were given in French. We assessed participants' skills with a touchscreen computer.

Children were presented with a sequence of colored pictures named by the experimenter as they appeared on the screen. Items were animals (e.g., a bee) or food (e.g., an orange) and were displayed for 2 seconds (see Figure 2). At the end of each sequence, children were asked to remember and sort the items they had seen and recite the list in the correctly sorted order. The number of items in the list started from two and increased every two trials to tax the working memory progressively. In the "1-list" task, pictures were only animals, and children had to sort them from smallest to largest. In the "2-list" task, pictures were animals and food; children had to sort food from smallest to largest, followed by animals from smallest to largest.



Question: Tell me the food from smallest to biggest and then the animals from smallest to biggest.

Correct response: Cherry, Rabbit and Tiger.

Figure 2: Example of the "2-list" sorting task with three items. Children must remember all items and sort the fruit from smallest to largest, followed by the animals from smallest to largest.

Flexibility

The flexibility task was a French adaptation of the NIH Toolbox Battery's Dimensional Change Card Sort (DCCS) on OpenSesame. The tasks assess children's cognitive flexibility by measuring their ability to switch between two different item sorting rules. Children were told that they would play the color and the shape games that each have their own rules. They were told that in the color game, they had to match items with the same color and that in the shape game, they had to match items with the same shape. Children were then shown two items at the bottom of the screen (e.g., a blue rabbit and a red boat). They had to sort ambivalent (e.g., red rabbits and blue boats) stimuli that appeared at the top of the screen (see Figure 3).

This modified version comprised four phases: familiarization, pre-switch, post-switch, and mixed. In the familiarization phase, children were instructed to follow a single rule (i.e., shape or color, counterbalanced across participants) to choose one of two target stimuli. For example, if they were instructed to follow the color rule when a red rabbit appeared at the top of the screen, the correct answer was the red boat at the bottom. The two target pictures were left in view during the whole phase. The children did two trials with each rule. The goal of this phase was to ensure that children understood the instructions clearly. No feedback was provided to the children.

In the following phases, the target items stayed displayed at the bottom of the screen. A fixation dot appeared for 1000 ms between each test item that appeared at the top of the screen, and a word appeared after the cross, indicating the rule to be followed. The rule was announced aloud by the experimenter. The computer's touch screen was used to record participants' responses, and test items disappeared after the child's response. Children had 10 000 ms to answer before the next trial appeared.

In the pre-switch phase, children saw five trials with one rule. In the post-switch phase, they saw five trials with the other rule. Rule order was balanced across participants.

In the mixed phase, children saw 30 trials, including 24 "frequent" and 6 "infrequent" trials presented in a pseudorandom order (with two to five frequent trials preceding each infrequent trial). For example, when the frequent rule was shaped, children would be instructed to sort by shape, but for six trials, they would be instructed to sort by color. Their ability to switch to the infrequent rule and back to the frequent rule allowed flexibility assessment.

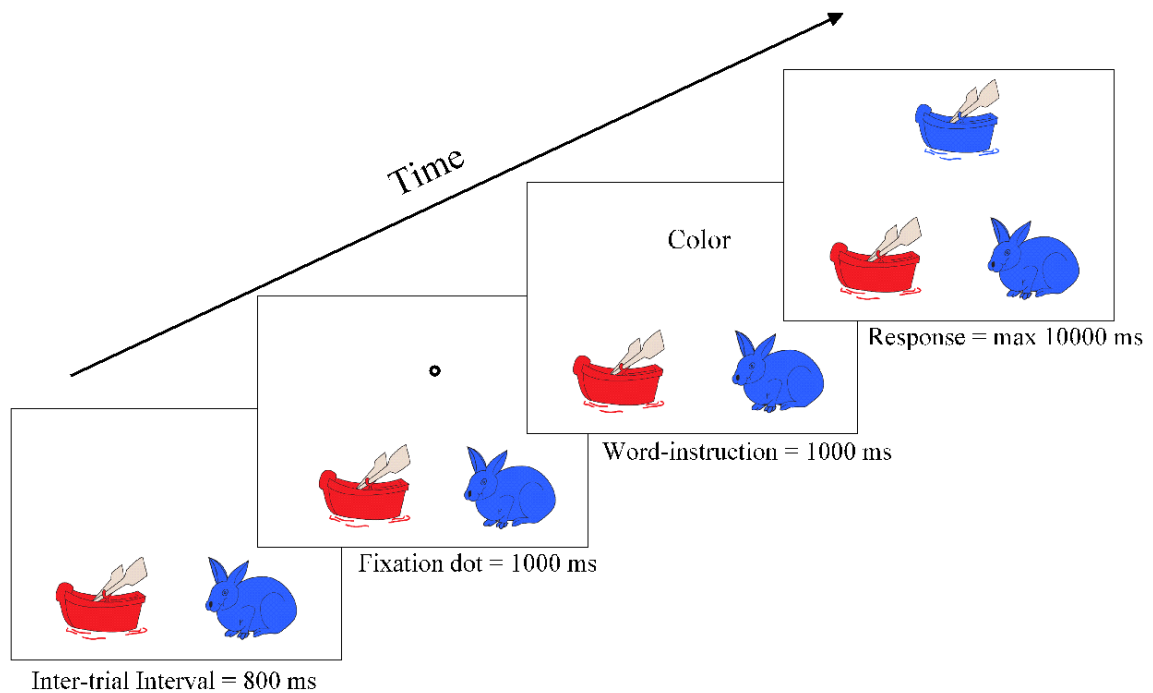


Figure 3: Example of one trial of the Dimension Change Card Sort (DCCS) task. This task requires children to sort cards by following either the shape or the color rule.

Inhibition

The inhibition task was a computerized French adaptation of the Real Animal Size Test (Catale & Meulemans, 2009), a nonalphabetical Stroop task adapted for young children. The task was run on OpenSesame. The animal Stroop task shows animals real life little or big animals (i.e., a butterfly, a bird, an elephant, a horse) in little or big size on the screen. Congruent trials are when real-life size is congruent with screen size. Congruent trials were: big elephants, big horses, small birds, and small butterflies. Incongruent trials were small elephants, small horses, big birds, and big butterflies. Children had to answer about animals' real-life size by pressing a key on the screen (right key for big, left key for small) (see Figure 4).

Before the task, we ensured that the children knew the animals by asking them to give their names. The task had three phases: control, training, and testing. In the control phase, items were all presented on screen in the same medium size. Children saw twelve trials and answered about the animals' real-life size. In the training phase, the animal's on-screen size changed: all animals were intermittently big or small on screen. Children were asked to answer about animals' real-life size. After a short practice (eight trials) where feedback was given, children saw sixteen trials (i.e., each animal in each size twice) without feedback, with

a limited time to answer (10 000 ms). A fixation dot was presented for 300 ms before every stimulus. In the testing phase, children saw thirty-two trials without feedback and were instructed to answer as quickly as possible.

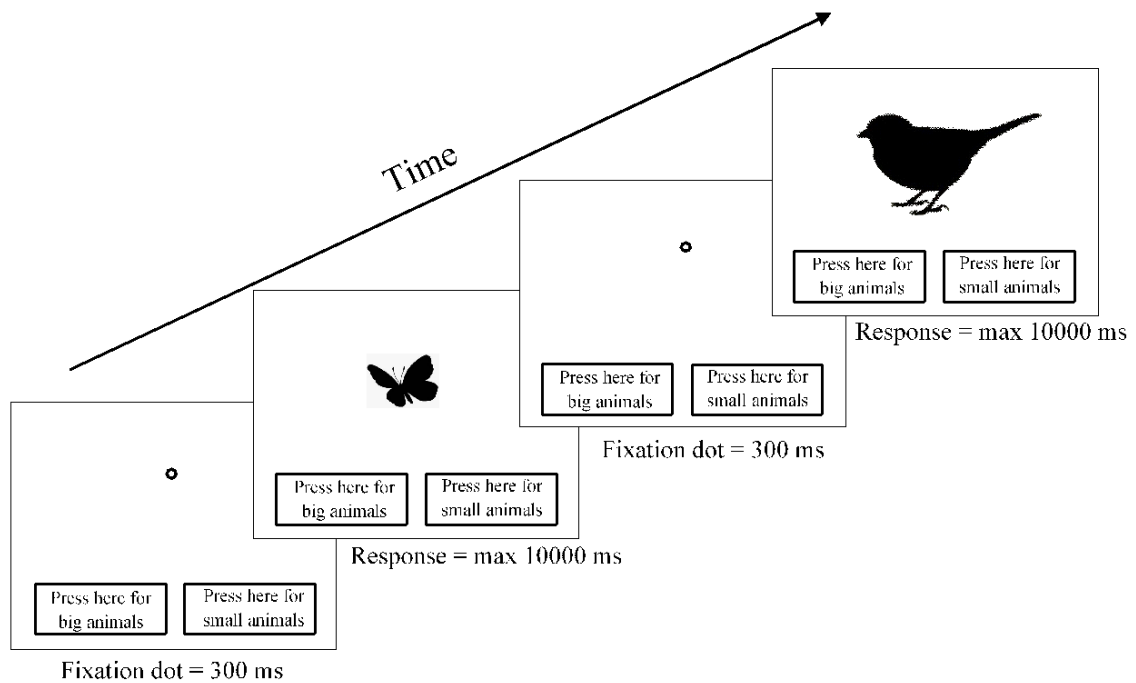


Figure 4: Example of two trials for the animal Stroop.

Visual-spatial reasoning

To assess visual-spatial reasoning we used the cube task from the fourth edition of the Wechsler Intelligence Scale for Children (*WISC-IV - Échelle d'intelligence de Wechsler pour enfants et adolescents - 4ème édition*, Wechsler, 2005). In this test, children had nine red and white cubes: each cub had two red sides, two white sides, and two red and white sides (see Figure X). Children had to place the cubes on a grid to replicate a given figure in a given time. Trials difficulty increased throughout the task: children started with four cube designs to replicate before moving on to nine cube designs (see Figure 5). The time given to complete a design varied as a function of the difficulty of the trial. The task ended when children failed to complete three successive designs in the given time.

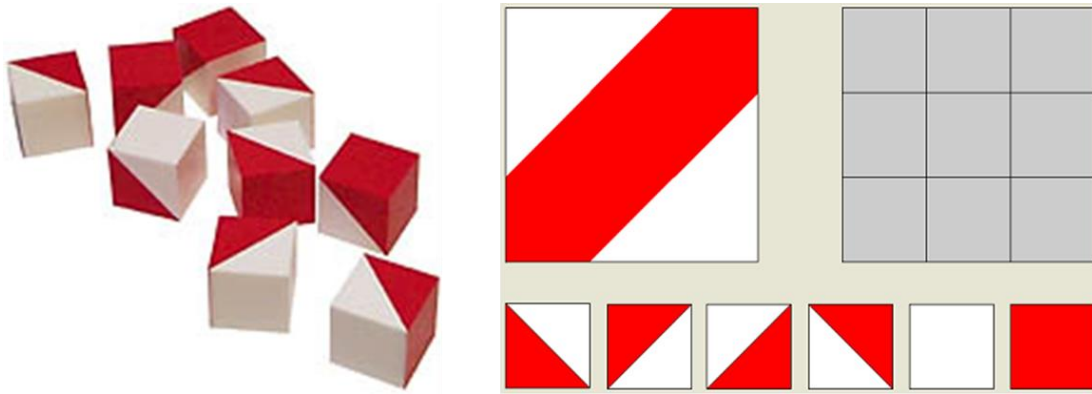


Figure 5. Cubes and example cards for the cube task from the Wisc-IV.

Procedure

Children were tested in a quiet room at their school. Cognitive assessments were made in two twenty-minute sessions after the generalization task session. Cognitive assessment and generalization sessions were one week to six months apart. The cognitive assessment tasks were shown in random order.

Data analysis

Vocabulary

EVIP scores were computed manually following the test's scoring chart. Standardized scores were not used.

Working memory

The working memory scores (List Sorting) were computed manually. Each correctly given sequence was accorded one point.

Flexibility

Reaction times and accuracy were recorded by OpenSesame. Incorrect responses were discarded from the reaction time analysis. All the RT inferior to 100 ms and superior to 10 000 ms or two deviation standards away from the mean were considered outliers and discarded from the analysis.

The flexibility score was using a two-vector method that incorporated both accuracy and, for participants who maintained a high level of accuracy (at least 80% correct), reaction time (Zelazo et al., 2013). This two-vector method was composed of an Accuracy score and an RT score. The Accuracy score was computed for all participants. The sum of the correct responses in the pre-switch (5 trials), the post-switch (5 trials,) and the mixed (30 trials) phase

was computed and times by 0.125 to obtain a score that ranged from 0 to 5. For children whose accuracy across all trials was less than 80% (i.e., Accuracy score < 4), the final score was equal to the Accuracy score. For children with 80% or higher accuracy, an RT score was also calculated based on each children's median RT on correct infrequent trials on the mixed block. Following the NIH procedure (Zelazo et al., 2013), a log (base 10) transformation was applied to each participant's median RT score, creating a more normal distribution of scores. Moreover, all the median RTs between 100 and 500 ms were set equal to 500 ms, and median RTs between 3000 ms and 10 000 ms were set equal to 3000 ms. Like the accuracy score, the RT score ranged from 0 to 5. Log values were algebraically rescaled with the following formula such that smaller RT log values were at the upper end of the 0–5 range, whereas larger RT log values were at the lower end. Once the rescaled RT scores were obtained, they were added to the Accuracy scores for participants who achieved the accuracy criterion of 80% or better.

$$\text{RT score} = 5 - 5 * \log \text{RT} - \log 500 \log 3000 - 500$$

Inhibition

Reaction times and accuracy were recorded by OpenSesame. Incorrect responses were discarded from the reaction time analysis. All the RT inferior to 100 ms and superior to 10 000 ms or two deviation standards away from the mean were considered outliers and discarded from the analysis.

We calculated an interference score by subtracting the mean proportion of correct answers in the incongruent condition from that of correct answers in the congruent condition.

Results

We described the cognitive assessment data for both tasks as a function of children's achievement level before testing for correlations between cognitive development and the percentage of correct taxonomic answers.

Relational noun generalization

Achievement profiles

Children's level of achievement was classified as follows. In relational noun generalization tasks, children who scored 12 correct answers or more (over 14 trials) were classified as high achieving, children who scored between 7 and 11 correct answers were

considered at chance, and children who scored less than 7 correct answers were classified as low achieving.

Data description as a function of achievement profiles

For each of the four cognitive assessment tests, we ran an ANOVA on the test's score with achievement level (low, chance, high) as a between factor.

Results revealed that significant effects for working memory $F(2,73) = 4.90, p < .01, \eta^2 = .12$, and cognitive flexibility $F(2,73) = 6.26, p < .01, \eta^2 = .15$. Post hoc Tukey comparisons were run for both effects. Working memory scores from high achieving children were significantly higher than those from children at chance in the generalization task ($M_{\text{High}} = 7.11; M_{\text{Chance}} = 5.02; M_{\text{Low}} = 6.24$; High-Chance: $p_{\text{Tukey}} < .01$). Cognitive flexibility scores were significantly higher for high and low achieving children compared to children at chance in the generalization task ($M_{\text{High}} = 5.33; M_{\text{Chance}} = 5.27; M_{\text{Low}} = 4.33$; High-Chance: $p_{\text{Tukey}} < .05$; Low-Chance: $p_{\text{Tukey}} < .05$).

Vocabulary and inhibition scores were not significantly different across achievement levels.

Cognitive mechanisms underlying novel relation noun generalization

To assess whether the cognitive functions evaluated were predictive of relational noun generalization performances, we ran Pearson correlations between participants' percentage of correct generalization and each cognitive assessment score. The correlations revealed that none of the cognitive assessment scores were correlated with the proportion of correct generalizations.

Table 1: Pearson correlations between children's percentage of correct relational noun generalizations and cognitive assessment scores.

	Generalization	Vocabulary	Working Memory	Flexibility	Inhibition
Vocabulary (EVIP)	$r = 0.148$ $p = 0.263$	-	-	-	-
Working Memory (List sorting, NIH Toolbox)	$r = 0.071$ $p = 0.544$	$r = 0.540$ $p < .001$	-	-	-
Flexibility (DCCS, NIH Toolbox)	$r = 0.136$ $p = 0.240$	$r = 0.294$ $p < 0.05$	$r = 0.351$ $p < 0.01$	-	-
Inhibition (Animal Stroop interference score)	$r = 0.063$ $p = 0.586$	$r = -0.297$ $p < 0.05$	$r = -0.077$ $p = 0.506$	$r = -0.062$ $p < 0.593$	-

Noun generalization

Achievement profiles

In noun generalization, children who scored 11 or more correct answers (over 14 trials) were classified as high achieving, children who scored between 5 and 10 correct answers were considered at chance, and children who scored 4 or less correct answers were classified as low achieving.

Data description as a function of achievement profiles

For each of the four cognitive assessment tests, we ran an ANOVA on the test's score with achievement level (low, chance, high) as a between factor.

Results revealed that significant effects for visual-spatial reasoning $F(2,159) = 4.90, p < .05, P2 = .05$, vocabulary $F(2,73) = 4.90, p < .01, P2 = .12$, and working memory $F(2,161) = 6.53, p < .01, P2 = .07$. Post hoc Tukey comparisons were run for all three effects. For all three cognitive processes (i.e., visual-spatial reasoning, vocabulary, and working memory), scores from high-achieving children were significantly higher than those from low-achieving

children (Visual-spatial reasoning, $p_{\text{Tukey}} < .05$; Vocabulary $p_{\text{Tukey}} < .05$; Working memory $p_{\text{Tukey}} < .001$) (see Table 2).

Table 2: Cognitive assessment means for visual-spatial reasoning, vocabulary, and working memory as a function of children's achievement level in a noun generalization.

Achievement level	Low	Chance	High
Visual-spatial reasoning (WISC, cubes)	21.8	22.0	26.1
Vocabulary (EVIP)	88.9	92.4	97.0
Working memory (List sorting, NIH Toolbox)	7.24	8.14	9.28

Cognitive mechanisms underlying novel noun generalization.

To assess whether the cognitive functions evaluated were predictive of relational noun generalization performances, we ran Pearson correlations between participants' percentage of correct generalization and each cognitive assessment score. The correlations revealed positive correlations between the percentage of correct generalizations and visual-spatial reasoning, vocabulary, and working memory (see Table 3).

We used general linear models with children's generalization scores (percentage of taxonomic generalizations) as outcomes and cognitive assessment scores as predictors, controlling for age (Table 4). The results did not reveal one cognitive assessment level as predictive of generalization performances.

Table 3: Pearson correlations between children's percentage of correct noun generalizations and cognitive assessment scores.

	Generalization	Visual-spatial reasoning	Vocabulary	Working memory	Inhibition
Generalization (%)	-	-	-	-	-
Visual-spatial reasoning (WISC, cubes)	$r = 0.235$ $p < .01$	-	-	-	-
Vocabulary (EVIP)	$r = 0.215$ $p < .01$	$r = 0.495$ $p < .001$	-	-	-
Working memory (List sorting, NIH Toolbox)	$r = 0.239$ $p < .01$	$r = 0.522$ $p < 0.001$	$r = 0.468$ $p < 0.001$	-	-
Inhibition (Animal Stroop interference score)	$r = -0.041$ $p = 0.605$	$r = 0.010$ $p = 0.0895$	$r = 0.035$ $p = 0.658$	$r = -0.005$ $p = 0.953$	-

Table 4: The goodness of fit of the generalized linear models with children's generalization scores as the outcome and working memory, spatial intelligence, and vocabulary scores as predictors.

The model controls for age.

Outcomes	Model	Predictors	Df	AIC	R ²	Adjusted R ²	p-value
Percentage of correct answers	M0	Age	1	1563	.083	.151	
	M1	Age + WM	1	1565	.080	.348	0.44
	M2	Age + WM + Spatial intelligence	1	1565	.082	.383	0.27
	M3	Age + WM + Spatial intelligence + Vocabulary	1	1567	.077	.399	0.62

Discussion

In the present work, we investigated which cognitive processes may support novel word generalization in comparison learning settings. Most studies in the domain infer cognitive processes from children's generalization decisions as a function of task demands (Gentner & Namy, 1999; Thibaut & Witt, 2017). Here, we took an individual difference assessment approach (Lagarrigue & Thibaut, 2020; Simms et al., 2018). We tested children's levels of executive function, spatial intelligence, and existing knowledge to evaluate how predictive of children's generalization behavior these processes may be.

Our main result was that spatial intelligence, existing knowledge, and working memory scores were correlated to novel noun generalization performances. This is a first argument to support our suggestion that information processing processes and existing knowledge are involved in novel noun generalization. We consider novel noun generalization tasks an example of concept building and use. Therefore, this first result supports the suggestion that concept building is a process in which the environment's stimuli and prior acquired knowledge interact.

Our results also revealed that higher-achieving children in word generalization had higher scores in multiple cognitive assessments. Regarding relational noun generalization, the higher-achieving children had significantly higher working memory and cognitive flexibility scores than other children. Regarding noun generalization, the high-achieving children had significantly higher spatial intelligence, working memory, and knowledge than other children. These results reveal that there is most likely a link between children's information processing mechanisms, knowledge, and word generalization.

However, our finding is not robust. Indeed, none of the cognitive processes positively correlated to generalization performances were predictive of generalization performances when tested in general linear regression models. The cognitive processes evaluated did not correlate with relational noun generalization performances in the first study either.

In the case of novel relational noun generalization, we suggest that the absence of results is due to the sample size. We tested seventy-six children, whereas in novel noun cases, we tested a total of one-hundred and sixty-two children, and Lagarrigue et al. (2022) tested one-hundred and twenty-one children. Another hypothesis may be the variation in children's age. Indeed, in novel relational noun cases, children are all five years old, whereas in noun generalization and in Simms et al. (2018), study children's ages varied between 5 and 8 years

old and 5 and 11 years old, respectively. In any case, the results are not highly consistent. This may signify that the link between information processing processes, knowledge, and generalization performance is more complex than previously hypothesized. For example, information processing capacities may mediate a cognitive process not yet studied that is itself predictive of generalization performances.

Another limit found here is the sensitivity of the inhibition task used, the animal Stroop. Even though the hypothesis about inhibition refraining automatic generalization choices of attractive distractors was strongly supported theoretically, inhibition was never correlated with generalization performance. However, this cannot refute the hypothesis completely. Indeed, inhibition scores reveal to have minimal variability across children. All children succeeded well and met a ceiling effect. A different way of assessing inhibition in young children is needed to test this hypothesis.

To conclude, it seems likely that both executive functions and knowledge support novel word generalization. However, how each process operates and at which level of the comparison and generalization process is still an open question.

References

- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons : A Meta-Analytic Review. *Educational Psychologist, 48*(2), 87-113. <https://doi.org/10.1080/00461520.2013.775712>
- Augier, L., & Thibaut, J.-P. (2013). The benefits and costs of comparisons in a novel object categorization task : Interactions with development. *Psychonomic bulletin & review, 20*(6), 1126-1132.
- Baldwin, D. A. (1989). Priorities in children's expectations about object label reference : Form over color. *Child Development, 60*, 1291-1306.
- Blaye, A., & Bonthoux, F. (2001). Thematic and taxonomic relations in preschoolers : The development of flexibility in categorization choices. *British Journal of Developmental Psychology, 19*(3), 395-412. <https://doi.org/10.1348/026151001166173>
- Blaye, A., & Jacques, S. (2009). Categorical flexibility in preschoolers : Contributions of conceptual knowledge and executive control. *Developmental Science, 12*(6), 863-873. <https://doi.org/10.1111/j.1467-7687.2009.00832.x>

- Catale, C., & Meulemans, T. (2009). The Real Animal Size Test (RAST) a new measure of inhibitory control for young children. *European Journal of Psychological Assessment*, 25(2), 83-91.
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64(1), 135-168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Dunn & Dunn. (1997). *Peabody Picture Vocabulary Test—Third Edition*. <https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft15145-000>
- Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure Mapping and Relational Language Support Children's Learning of Relational Categories : Structure Mapping and Relational Language. *Child Development*, 82(4), 1173-1188. <https://doi.org/10.1111/j.1467-8624.2011.01599.x>
- Gentner, D., & Namy, L. L. (1999). Comparison in the Development of Categories. *Cognitive Development*, 14(4), 487-513. [https://doi.org/10.1016/S0885-2014\(99\)00016-7](https://doi.org/10.1016/S0885-2014(99)00016-7)
- Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology*, 107(3), 280-290. <https://doi.org/10.1016/j.jecp.2010.04.017>
- Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting Long Distance : Semantic Distance in Analogical Reasoning Modulates Frontopolar Cortex Activity. *Cerebral Cortex*, 20(1), 70-76. <https://doi.org/10.1093/cercor/bhp081>
- Hammer, R., Bar-Hillel, A., Hertz, T., Weinshall, D., & Hochstein, S. (2008). Comparison processes in category learning : From theory to behavior. *Brain Research*, 1225, 102-118. <https://doi.org/10.1016/j.brainres.2008.04.079>
- Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning : The role of shape similarity in early acquisition. *Cognitive Development*, 9(1), 45-75. [https://doi.org/10.1016/0885-2014\(94\)90019-1](https://doi.org/10.1016/0885-2014(94)90019-1)
- Keil, F. C. (1992). *Concepts, Kinds, and Cognitive Development*. MIT Press.
- Lagarrigue, Y. (2022). *Apprentissage et généralisation de nouveaux concepts chez l'enfant par la comparaison de stimuli* [These de doctorat, Bourgogne Franche-Comté]. <https://www.theses.fr/2022UBFCK034>
- Lagarrigue, Y., & Thibaut, J.-P. (2020). *From two to many : The role of executive functions in young children's generalization of novel object names in a comparison design*.
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development*, 3(3), 299-321.

- Markman, A. B., & Gentner, D. (1993). Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, 25(4), 431-467. <https://doi.org/10.1006/cogp.1993.1011>
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, 92, 289-316. <https://doi.org/10.1037/0033-295X.92.3.289>
- Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears : Young children's use of comparison in category learning. *Journal of Experimental Psychology: General*, 131(1), 5.
- Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognition, Creier, Comportament/Cognition, Brain, Behavior*, 11(4).
- Peabody Picture Vocabulary Test—Third Edition*. (s. d.). Consulté 28 mars 2023, à l'adresse <https://psycnet.apa.org/doiLanding?doi=10.1037%2Ft15145-000>
- Price, G. F., & Sandhofer, C. M. (2021). One versus many : Multiple examples in word learning. *Journal of Experimental Child Psychology*, 209, 105173. <https://doi.org/10.1016/j.jecp.2021.105173>
- Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning : Insights from scene analogy problems. *Journal of Experimental Child Psychology*, 94(3), 249-273. <https://doi.org/10.1016/j.jecp.2006.02.002>
- Simms, N. K., Frausel, R. R., & Richland, L. E. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology*, 166, 160-177. <https://doi.org/10.1016/j.jecp.2017.08.005>
- Sloutsky, V. M. (2010). From perceptual categories to concepts : What develops? *Cognitive science*, 34(7), 1244-1286.
- Stansbury, E., Witt, A., & Thibaut, J.-P. (2022). How Stimuli Availability Effects Novel Noun Generalization in a Free-Choice Design. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44(44), 2789-2795.
- Thibaut, J.-P., French, R., & Vezneva, M. (2010). The development of analogy making in children : Cognitive load and executive functions. *Journal of Experimental Child Psychology*, 106(1), 1-19. <https://doi.org/10.1016/j.jecp.2010.01.001>
- Thibaut, J.-P., Stansbury, E., & Witt, A. (2018). Generalization of novel names for relations in comparison settings : The role of conceptual distance during learning and at test. *Livre/Conférence Proceedings of the 40th Annual Meeting of the Cognitive Science Society*, 1114-1119.

- Thibaut, J.-P., & Witt, A. (2015). Young children's learning of relational categories : Multiple comparisons and their cognitive constraints. *Frontiers in Psychology*, 6.
<https://doi.org/10.3389/fpsyg.2015.00643>
- Thibaut, J.-P., & Witt, A. (2017). Generalizing novel names in comparison settings : Role of conceptual distance during learning and at test. *CogSci 2017 : 39th Annual Meeting of the Cognitive Science Society*, 3314-3319. <https://hal-univ-bourgogne.archives-ouvertes.fr/hal-01573610>
- Ware, E. A. (2017). Individual and developmental differences in preschoolers' categorization biases and vocabulary across tasks. *Journal of Experimental Child Psychology*, 153, 35-56.
<https://doi.org/10.1016/j.jecp.2016.08.009>
- Wiebe, S. A., & Karbach, J. (2017). *Executive function : Development across the life span*. Routledge.
- WISC-IV - Échelle d'intelligence de Wechsler pour enfants et adolescents—4ème édition.
(s. d.). Pearson Clinical & Talent Assessment. Consulté 28 mars 2023, à l'adresse
<https://www.pearsonclinical.fr/wisc-iv-echelle-dintelligence-de-wechsler-pour-enfants-et-adolescents-quatrieme-edition>
- Zelazo et al. (2013). *INIH TOOLBOX COGNITION BATTERY (CB) : MEASURING EXECUTIVE FUNCTION AND ATTENTION*.
<https://srcd.onlinelibrary.wiley.com/doi/abs/10.1111/mono.12032>

**SECTION 3: DIFFERENT INSIGHTS INTO
GENERALIZATION WITH A DIFFERENT PARADIGM**

Section 3: Different insights into generalization with a different paradigm

Introduction

Children remarkably succeed in learning language from a limited number of stimuli associated with the words they learn (Bloom, 2002; Carey & Bartlett, 1978; Murphy, 2002). Understanding how children make sense of this paucity in novel word learning and generalization situations and how these tasks may favor accurate word generalization is crucial for understanding language development.

Generalization constraints and free-choice designs

A prominent view on word generalization is that children's attention is directed to particular stimuli dimensions that are then used as a bases for word extension (Jones & Smith, 1993; Markman, 1989; Waxman & Kosowski, 1990). How one's attention is directed and to which dimensions attention is directed are crucial questions in understanding what determines young children's word generalizations. It can be considered that multiple constraints are at play in directing attention, in an either top-down or bottom-up way, such as perceptual feature salience, instructions, prior knowledge in a perceptually grounded or lexicalized form, context etc. When these constraints are particularly strong, they bias children's generalization. This is the case for example with the shape bias where children generalize words towards targets with the same shape rather than the same color, size or texture (Kucker et al., 2019b; e.g., Landau et al., 1988c), which can be explained by the greater salience of shape over other dimensions, that directs children's attention in a bottom-up way (Smith et al., 2002). To succeed in generalizing words on a conceptual basis – which is what is required for a correct use of

words and successful communication – constraints that guide attention towards conceptual information must be stronger than those that orientate towards other dimensions.

Correct word generalization is consequently largely dependent of how learning and generalization are constrained by the stimuli, the context, prior knowledge etc. From this standing point we considered how traditionally used word generalization tasks constrain children's word extensions. An overwhelming proportion of word generalization studies consider forced-choice task designs (Augier & Thibaut, 2013; Childers, 2011; Gentner & Namy, 1999; Graham et al., 2010; Hammer & Diesendruck, 2005; Landau, 1994; Namy et al., 2007b; Thibaut et al., 2018). The forced choice design is well suited to study biases as they tell us what children select with the highest probability among a set of options. However, it appears that this type of design constrains word generalization in a restrictive way for studying children's conceptual development. Indeed, in forced-choice designs children generalization choices reveals which item they considered is related to the learning items in the best way. This is instructive of children's categorical decision in a given context. However, children may be choosing an item as the word extension target that they may not have considered a category member otherwise because they are instructed to choose at least one item: "which one is also a xx". Conversely, children may choose the best match while considering other items as plausible category members and word extension targets, without having the opportunity to select them, given the instructions. By limiting answers to a single item, forced choice designs say nothing of the scope of items children consider as category members and plausible word extension targets. For example (Smiley & Brown, 1979) showed that young children in a forced choice design could select and justify a taxonomic choice even when their first choice was a thematic choice.

To overcome this limit of forced-choice novel word generalization tasks, suggest studying novel noun generalization with free-choice designs. This design will allow us to portray a more realistic scope of items children consider plausible generalization targets. The empirical work that stemmed from this suggestion is presented and discussed in the following section.

Throughout our work with free choice designs our main aim was still to get a better understanding of *how* children come to a generalization decision. For investigating this question free-choice designs provide two interesting lines of investigation, namely manipulating generalization item availability and using a signal detection theory statistical analysis framework.

Free choice designs enable one to consider the role of generalization items in generalization decisions by allowing one to manipulate generalization item availability. In a free choice task the experimenter will ask “Is this also a Buxy?” entailing a yes-no answer from the child. The design may display all generalization item possibilities simultaneously (i.e. taxonomic answers and distractors) therefore asking the child “Are there any other Buxies?”, or the design may, like perceptual discrimination tasks, display each possible generalization item one-by-one in a sequential manner asking the child “Is this also a Buxy?” for each item. Consequently, comparing results in different generalization availability settings (i.e., whether all generalization items are presented simultaneously or sequentially) informs on the effect of the generalization items on the generalization decision.

Second, in free choice tasks generalization scores result from yes-no decisions and can therefore be analyzed in a signal detection framework. This framework gives one the possibility to consider more aspects of generalization behavior. Indeed, signal detection analyses consider both sensitivity (i.e. ability to discriminate the signal among noise) and response strategy (i.e., willingness to answer) by considering generalization choices as well as falls alarms, correct rejections and falls rejections that are most often left aside because they can't be considered in traditional accuracy or percentage of choice analyses.

Before proceeding we will briefly present how generalization availability and signal detection theory can be understood in the case of novel noun generalization.

Generalization item availability

It is commonly considered that comparison between items at learning is important for building the mental representation used for future generalization because comparison enhances pertinent dimensions for generalization. Surprisingly, other items or elements of context that may help or hinder building a viable representation for future generalization are rarely considered. For example, the influence of the task's generalization items is admitted while it is rarely studied. This has a priori flaws. First, comparison is considered one of the most powerful drivers of learning (Alfieri et al., 2013a; Gentner & Namy, 1999, 2004b; Graham et al., 2010) that is supported by specific brain mechanisms (Seger, 2008). If one considers comparison as a main process in learning it is unlikely that comparison between learning and generalization items or between generalization items themselves may play a role in children's generalization. This role has yet to be specified. Second, it is unlikely that children are completely unaware of generalization items while building their mental

representation. In generalization tasks that presents all items simultaneously it is highly unlikely that children ignore all generalization items until their mental representation is extracted from the learning items. Even in settings where generalization choices are produced after learning, items re-representation processes (Gentner & Colhoun, 2010b) are likely to intervene and adapted the representation. Moreover, our eye-tracking analyses reveal that attention to generalization items sets of early in generalization tasks which confirms that generalization items are likely involved in the representation building process. Noun generalization strategies were also revealed as elimination type strategies, were children match learning and generalization items and mentally compare these matches to one another in order to find the best match.

Our results on novel noun generalization strategies also support the suggestion that generalization items play a role in conceptualization during generalization. The strategy revealed by eye tracking data was an elimination type strategy. Children searched for a generalization solution by mapping learning items to each of the generalization items, eliminating matches until the best match is found. This hypothesis about children's strategy implies that children chose a best learning-generalization mapping and therefore compare learning-generalization mappings to each other to find the best match. Even if very few eye movements between items are revealed, it seems likely that the various possibilities play a certain role in children's generalization decisions.

Considering these observations, we suggest that generalization items are involved in the building of mental representations during word generalization. In the studies presented in this section, we manipulated generalization item availability, a factor with two variables: simultaneous and sequential availability. We suggest that multiple taxonomic generalization items, presented simultaneously, highly increase the amount of conceptual information available while the mental representation is being built, and will help constrain generalization towards conceptually related items. Reversely, when generalization items are presented sequentially, each generalization decision is made with only the learning item's information and the information given by the single generalization item. Thus, children should display more conceptual generalizations in simultaneous settings in which more input directs attention towards conceptual choices than they do in sequential settings.

Signal detection theory

Signal detection theory is a methodological approach for measuring performances that involve finding a “signal” in background noise (Macmillan & Creelman, 2005). While it was at first designed for perceptual tasks it has been used for many other and more complex cognitive tasks in which participants encounter stimuli of different types and must assign different responses to them. The suggestion that supports the theory is that a participant places a threshold, called decision criterion, on a continuous perceptual domain and categorizes stimuli as signal or noise according to their perceived position to the threshold.

One important advantage of signal detection analyses is that they consider more aspects of participants response behavior than traditional accuracy or proportion of answer analyses. Signal detection theory indexes are calculated from all of participants answers classified in four sets: participants hit rate (i.e., signals categorized as signals), falls alarms (i.e., not signal stimuli classified as stimuli), misses (i.e., signal classified as noise) and correct rejections (noise classified as noise). Two indexes are then calculated to describe both participants’ sensitivity to the signal (i.e., ability to discriminate the signal among the noise) and participants response strategy (i.e., willingness to answer).

In previous research the adaption of signal detection theory to conceptual development models led to decision bound models (Maddox & Ashby, 1993). These models’ assumption is that when categorizing an object, participants assign a response according to the region in perceptual space in which the object falls. Perceptual space is partitioned by a *decision bound* between competing regions, and the regions position to the bound determines the objects category. In the more widely admitted exemplar view however, category membership decisions are the product of the sum of perceived similarities between the signal and every exemplar already experienced and stored in memory. Our aim is not to take part in this debate about the most realistic conceptual model, and our work does not allow us to support one model more rather than the other. However, both models can be kept in mind considering we apply signal detection theory to novel noun generalization tasks, in free choice designs.

What is however necessary before considering the following studies is defining a framework in which to consider signal detection applied to novel noun generalization. Namely, we must consider what signal detection indexes (i.e., sensitivity and response strategy) may represent in novel noun generalization tasks.

Sensitivity is participants ability to distinguish signal from noise (Macmillan & Creelman, 2005). High sensitivity is participants good ability to identify the signals leading to hits and correct rejections, and low sensitivity is a poorer ability to discriminate, leading to missed signals or false alarms. In novel noun generalization, sensitivity represents participant's ability to distinguish taxonomically related items among distractors. We consider, as presented in the General introduction that participants ability to distinguish taxonomically related items however depends on the item's similarity to the category's representation (i.e., the concept) stored in memory. We therefore suggest that high sensitivity results from a dense and detailed representation while lower sensitivity results from a sparse or fuzzy representation. In other words, the better the representation's quality for identifying category members, the more pertinent conceptual information it contains, the higher participants sensitivity will be.

Response bias is participants tendency to respond with some degree of favoritism or their tendency to tilt towards one response rather than another (Macmillan & Creelman, 2005). Response bias is also considered as participants willingness to respond, for example when answering implies a certain level of risk. In novel noun generalization task, there is no – or an extremely low – risk in answering. However, one can consider that willingness to answer can be considered, not as a result of a risk, but as a result of the presence of a categorization rule that may be followed more or less willingly. We suggest that in novel noun generalization, response strategy represents participants inclination or willingness to follow the taxonomic rule when extending the novel nouns. Response strategy is also called response bias and varies from a conservative bias (i.e., where participants are unwilling to answer leading to many misses), to a liberal bias (i.e., where participants answer widely leading to false alarms). In novel noun generalization, in which participants are free to decide about the rule they choose to follow when generalizing the word, we suggest that response strategy reflects participants certainty about the categorization rule to follow. A conservative bias will then reflect a high level of uncertainty about the rule they chose to generalize by, a neutral bias will reflect confidence in the generalization rule chosen, while a liberal bias will reflect a biased generalization process or an “overconfident” generalization process.

In order to calculate these biases the experimenter records participants number of hits (H), correct rejections (CR), misses (MS), and false alarms (FA). Hits and correct rejections are correct categorization answers (e.g., hits: chosen taxonomic items, correct rejections: rejected distractors), whereas misses and false alarms are wrong categorization answers

(misses: unchosen taxonomic items; false alarms: chosen distractors). Sensitivity and response bias indexes were computed as follows, with formulas adapted to small signal sets taken from Rioux et al., (2018) (used in Morin-Audebrand et al., 2012).

- Sensitivity index = $\log[(H + 0.5)/(MS + 0.5)] - \log[(FA + 0.5)/(CR + 0.5)]$
- Response bias index = $-\log[(H + FA + 0.5)/(MS + CR + 0.5)]$

The present studies

In the study presented in Chapter 1, we tested our main hypothesis: generalization items influence mental representations built in novel noun generalization tasks and thus generalization performances. In this case, an effect of generalization availability on generalization performances should be revealed, showing better generalization performances in simultaneous availability settings. Three- and four-year old children were tested in a free choice novel noun generalization task in which learning items were perceptually dissimilar, and six generalization items were presented: two items from the learning items' near superordinate category (i.e., near superordinate items), two items from the learning items' distant superordinate category (i.e., distant superordinate items), and two items thematically related to the learning items.

Building on this first studies results, our second study – presented in Chapter 2 - aimed at giving a more complete description of novel noun generalization in free choice designs, and of the effect of generalization item's availability. The first study had revealed that four-year-old is the youngest age at which the task could be considered, and our aim was to observe the developmental course of results from that age with two other age groups. Thus, we tested four-, five-, and six- year-old children in a free choice novel noun generalization task. We manipulated conceptual distance between learning items (close and far) and learning items were dissimilar or similar. As already mentioned, our aim was to analyze a large scope of what items children consider as possible generalization targets and how this scope of possibilities is affected by generalization availability. To give a complete description as possible of this scope of considered possibilities, we included taxonomically related items from both near and distant superordinate categories, and all types of distractors commonly used in novel noun generalization tasks. Thus, at generalization children saw 10 generalization items: two near superordinate items, two distant superordinate items, two items thematically related to the learning items, two perceptually related items and two unrelated

items. Another group of five-year-old children were also tested in a no-comparison setting with 10 generalization items to set a base line for generalization in a no-comparison setting in a free choice design which had never been done.

Very low level of choices in sequential availability while choices in simultaneous. Children can choose but don't want to choose. Considering same basic level items. And choice behavior for same taxo level items.

The rather extensive description of children's generalization provided by this second study revealed surprisingly low percentages of choices for taxonomically related items – even near superordinate items- especially when generalization item availability was sequential. Our reasoning was that forced choice designs with children from these age groups, and in a certain extent simultaneous settings, reveal that children can generalize novel words to near superordinate items, but results from our second study revealed that they were preferring to not generalize to these items. We suggest that, when given the choice as they are in free-choice designs, children did not find the near superordinate items a satisfying generalization target but may find a same basic category level item a good match. In another study presented in Chapter 3 we included two generalization items from the same basic level as the learning items. Children saw 12 generalization items, the same as in the previous study and two same basic category level items for each category. Five-year-old children were tested with similar learning items.

References

- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons : A Meta-Analytic Review. *Educational Psychologist*, 48(2), 87-113. <https://doi.org/10.1080/00461520.2013.775712>
- Augier, L., & Thibaut, J.-P. (2013). The benefits and costs of comparisons in a novel object categorization task : Interactions with development. *Psychonomic bulletin & review*, 20(6), 1126-1132.
- Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
- Carey, S., & Bartlett, E. (1978). *Acquiring a single new word*.
- Childers, J. B. (2011). Attention to multiple events helps two-and-a-half-year-olds extend new verbs. *First Language*, 31(1), 3-22. <https://doi.org/10.1177/0142723710361825>

- Gentner, D., & Colhoun, J. (2010). Analogical Processes in Human Thinking and Learning. In B. Glatzeder, V. Goel, & A. Müller (Éds.), *Towards a Theory of Thinking* (p. 35-48). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03129-8_3
- Gentner, D., & Namy, L. L. (1999). Comparison in the Development of Categories. *Cognitive Development, 14*(4), 487-513. [https://doi.org/10.1016/S0885-2014\(99\)00016-7](https://doi.org/10.1016/S0885-2014(99)00016-7)
- Gentner, D., & Namy, L. L. (2004). The Role of Comparison in Children's Early Word Learning. In *Weaving a lexicon* (p. 533-568). Boston Review. <https://doi.org/10.7551/mitpress/7185.001.0001>
- Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107*(3), 280-290. <https://doi.org/10.1016/j.jecp.2010.04.017>
- Hammer, R., & Diesendruck, G. (2005). The role of dimensional distinctiveness in children's and adults' artifact categorization. *Psychological Science, 16*(2), 137-144.
- Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8*(2), 113-139.
- Kucker, S. C., Samuelson, L., Perry, L., Yoshida, H., Colunga, E., Lorenz, M., & Smith, L. (2019). Reproducibility and a Unifying Explanation : Lessons from the Shape Bias. *Infant behavior & development, 54*, 156-165. <https://doi.org/10.1016/j.infbeh.2018.09.011>
- Landau, B. (1994). Object shape, object name, and object kind : Representation and development. *D. Medin (Ed.). Psychology of Learning and Motivation: Advances in Research and Theory, 31*, 253-304.
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development, 3*(3), 299-321.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory : A user's guide* (2nd ed). Lawrence Erlbaum Associates.
- Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics, 53*(1), 49-70. <https://doi.org/10.3758/BF03211715>
- Markman, E. M. (1989). *Categorization and naming in children : Problems of induction*. Mit Press.

- Morin-Audebrand, L., Mojet, J., Chabanet, C., Issanchou, S., Møller, P., Köster, E., & Sulmont-Rossé, C. (2012). The role of novelty detection in food memory. *Acta Psychologica, 139*(1), 233-238. <https://doi.org/10.1016/j.actpsy.2011.10.003>
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognition, Creier, Comportament/Cognition, Brain, Behavior, 11*(4).
- Rioux, C., Leglaye, L., & Lafraire, J. (2018). Inductive reasoning, food neophobia, and domain-specificity in preschoolers. *Cognitive Development, 47*, 124-132. <https://doi.org/10.1016/j.cogdev.2018.05.001>
- Seger, C. A. (2008). How do the basal ganglia contribute to categorization? Their role in generalization, response selection, and learning via feedback. *Neuroscience and biobehavioral reviews, 32*(2), 265-278. <https://doi.org/10.1016/j.neubiorev.2007.07.010>
- Smiley, S. S., & Brown, A. L. (1979). Conceptual preference for thematic or taxonomic relations : A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology, 28*(2), 249-257.
- Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological science, 13*(1), 13-19.
- Thibaut, J.-P., Stansbury, E., & Witt, A. (2018). Generalization of novel names for relations in comparison settings : The role of conceptual distance during learning and at test. *Livre/Conférence Proceedings of the 40th Annual Meeting of the Cognitive Science Society*, 1114-1119.
- Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations : Toddlers' and preschoolers' word-learning biases. *Child development, 61*(5), 1461-1473.

Chapter 1 - Generalization of novel object names in comparison contexts in a yes-no paradigm by young children. When the rate of stimulus presentation matters.

The following chapter contains the proceedings paper: Stansbury E., Witt A., Thibaut J-P., (2020). Generalization of Novel Object Names in Comparison Contexts in a yes-no paradigm by young children. When the rate of stimulus presentation matters. *In Proceedings of the 42nd Annual Meeting of the Cognitive Science Society (Eds.)*, (pp. 2288 - 2294).

Note: Additions made to the paper for the clarity of the thesis are in boxes. The terminology used was unified across the thesis, changes to this paper were: Near = Near superordinate, Distant = Distant superordinate, Transfer Mode = Availability, Transfer items = Generalization items.

Generalization of Novel Object Names in Comparison Contexts in a yes-no paradigm by young children. When the rate of stimulus presentation matters.

Eleanor Stansbury (eleanor.stansbury@u-bourgogne.fr)

Arnaud Witt (arnaud.witt@u-bourgogne.fr)

Jean-Pierre Thibaut (jean-pierre.thibaut@u-bourgogne.fr)

LEAD-CNRS, UMR 5022, Université de Bourgogne Franche-Comté,

Pôle AAFE – Esplanade Erasme, 21065 Dijon, France

Abstract

A common result in novel word generalization tasks is that comparison settings (i.e., several stimuli introduced simultaneously) favor conceptualization and generalization. We hypothesized that typical comparison forced-choice designs between a lure and a targeted item might have constrained children's choices. Here we used a "yes-no" free choice design with 3- and 4-year-old children, and manipulated the generalization item's availability, either simultaneous or sequential. We manipulated the semantic distance between training and generalization items. Results showed that simultaneous, rather than sequential, presentations in the generalization phase led to more taxonomic generalizations in four-year olds. Results are discussed in terms of the constraints that both types of presentation bring into the task.

Keywords: novel name; comparisons; generalization; forced-choice; free-choice.

Introduction

When they learn to categorize and name novel objects, children have to capture which dimensions are important to define the corresponding concept (Murphy, 2002). One difficulty is that, in some cases, perceptual similarities (e.g., objects from different categories displaying the same texture and/or the same color) or differences are more salient than variations along the relevant features. Semantic similarities (e.g., thematic similarities) can also be misleading. It can therefore be challenging for young children to ignore salient but irrelevant perceptual or semantic similarities and generalize according to less salient but deeper, conceptually based properties (Augier & Thibaut, 2013, Gentner & Namy, 1999). In this context, understanding which presentation format(s) and which strategies lead to conceptually driven generalization is an important topic for cognitive sciences.

There is now considerable evidence that the opportunity to compare stimuli during learning highlights nonobvious shared properties and favors conceptually based categorization and novel word generalization more than the classical single learning exemplar situations. However, little is known of the dynamics of comparisons in word generalization tasks. Which items are compared? In which order? If children benefit from comparisons between learning items, do they also benefit from comparisons between learning-items and generalization-items? This study aims to assess the impact of comparisons made between learning-items and generalization-items in a traditional name generalization task by manipulating the generalization items' availability, either favoring one-by-one comparisons between

the learning items and each generalization item (sequential availability), or allowing to compare learning items to the whole set of generalization-items presented together (simultaneous availability).

Comparisons favor taxonomic generalization

Most of the existing comparison studies with children manipulate familiar objects, relations or situations and to the best of our knowledge, in all previous studies, children were asked to extend the novel name in a forced-choice design (Christie & Gentner, 2010). In a typical design, the learning items are perceptually similar (i.e. they display the same shape). The child has to generalize the learning items name to one of the generalization items. One of the generalization items, the taxonomic item, is taxonomically related to the learning items and is rated as less perceptually similar to them. Another generalization item is a perceptually similar lure that is conceptually unrelated to the learning items but perceptually similar to them, or to one of them (e.g., Gentner & Namy, 1999). The lure can also be thematically related to the learning items (e.g., a fork for two food items). Many studies have shown that comparison situations lead to more conceptually based generalizations than no-comparison situations.

Many studies have investigated which conditions help children extract a taxonomic relation during the learning phase and generalize it to the correct generalization item. Overall, we know that the presence of a common name invites comparison (Gentner & Namy, 1999) and contributes to conceptually based generalization. We also know that increasing the number of compared items does not automatically lead to more conceptually-based generalization, especially in younger children, most likely because of the cognitive costs induced by multiple comparisons (Augier & Thibaut, 2013; Thibaut & Witt, 2015) and that semantic distance might also matter. Thibaut and Witt (2017) manipulated the semantic distance between learning items (e.g., two bracelets versus a bracelet and a watch), and the semantic distance between the learning items and the generalization items (e.g., a jewel, near superordinate distance, versus a bow tie, distant superordinate distance). Six-year-olds made more distant generalization in the far learning condition than in the close learning condition. Children also made less taxonomic choices in the distant generalization condition compared to the near superordinate

generalization condition, confirming that increasing the generalization distance between learning and generalization items makes the task harder for young children, especially when the compared learning items were very close (e.g. two apples).

Comparison Strategies

As mentioned above, most comparison designs are forced-choice designs, in which children first see the learning items together with their name. Immediately after, they are shown two (or more) options, one being a correct conceptually-based generalization item, the other being incorrect (a perceptual or a thematic lure).

Forced-choice tasks are reasoning tasks in the sense that children have to choose the most plausible answer given the available evidence. However, children could choose an item that is plausibly related to the standards but that they would not select as an item of the same category if they were not forced to choose one. Moreover, a selection of a most plausible option does not entail that participants would not accept the other option as a member of the category if they were given the opportunity to select it. In a forced-choice design, Smiley and Brown (1979) showed that young children could select and justify a taxonomic choice even when their first choice was a thematic choice. Forced-choice designs are well suited to study what commonalities children are able to generalize. Free-choice designs give other insights on the items children would *really* generalize as members of a category.

Goals of the present experiment

In the following experiment, we use a comparison design and contrast two theories of generalization. The first derives from most comparison studies and suggests that the main determinants of conceptualization are the comparisons between learning items, during the learning phase. This view conceives learning as the product of the comparisons taking place with the learning items and gives no status to comparison during generalization. Another possibility is that conceptualization is also determined by other comparisons between other available items, for example items available in the noun generalization phase.

In contrast with forced-choice designs, we used a free-choice task in which we asked participants to select the items they thought would also hold the same name, with no additional constraint. Our main goal was to contrast a simultaneous with a sequential availability. In real life, children often see generalization items one by one, in a sequential availability, in which they encounter generalization items on successive occasions. Thus, it is difficult to make systematic comparisons between learning items and different generalization items or to compare generalization items one with the others. In comparison, in the simultaneous availability children saw all the generalization items simultaneously, a mode which is more similar to presentations of various stimuli in books. The simultaneous condition, compared with the sequential

condition, allow for additional types of comparisons, that is comparisons between learning items and all the generalization items and, second, between generalization items themselves, so that children might decide which stimuli are the most adequate generalization items. How would these additional comparisons contribute to generalization? One possibility would be that comparisons in the simultaneous mode would allow children to align distant superordinate generalization items with near superordinate generalization items and extend their generalization to more distant superordinate items. Both comparisons between learning items and generalization items, and between generalization items themselves also contribute to highlight the irrelevance of the distractor thematic relation and should contribute to exclude these thematically related objects from the chosen generalization items.

Along this latter theoretical perspective, our first prediction is that the simultaneous mode may lead to better taxonomic generalization performance because these comparisons will contribute to highlight the relevant properties. This means that participants should select more taxonomic choices but also less non taxonomic lures in the simultaneous condition.

However, this factor, availability, might interact with the semantic distance and age. As for the semantic distance between learning and generalization items, we mentioned above that it influences the scope of children's conceptual generalization. We predict more near superordinate generalization item (i.e., items that belong to the close categories) choices than distant superordinate generalization item (i.e. items from remote categories) choices in both the simultaneous and the sequential modes, because it is easier to understand that near superordinate generalization items are of the same kind as the learning items. However, we also predict that availability and semantic distance might interact. Our conceptual framework predicts that difference between near superordinate and distant superordinate generalization items should be smaller in the simultaneous mode because participants will be able to compare near superordinate generalization items to distant generalization items in this mode. These comparisons between near superordinate and distant superordinate generalization items may help participants to extend their generalization whereas such an alignment process is impossible in the sequential mode.

We also predict, for the same alignment reasons, that a learning condition including both semantically close and far learning items should lead to better results in both conditions than learning conditions made up with either of close learning items or of far learning items. A mixture of both might contribute to better conceptualize the scope of the category.

We also make predictions regarding interactions between age and the other factors. Generalization is a cognitively demanding task and previous studies have shown that young children can be overwhelmed by excessive quantities of information in generalization tasks (Augier & Thibaut, 2013; Stansbury, Witt, & Thibaut, 2018; Thibaut & Witt, 2015). We hypothesize that the sequential mode should be more cognitively demanding and thus predicts that the difference

between younger and older children should be larger in the sequential than in the simultaneous conditions.

By contrast, if conceptualization and generalization is mainly a matter of comparisons involving the learning items, then later comparisons between generalization items should not influence children's generalization. This predicts equivalent generalization results between *simultaneous* and *sequential* availability. This view does not make any specific suggestions regarding the generalization items semantic distance, or between generalization and age.

Addition to the paper:

The free choice design used here also allows us to use a signal detection theory analysis. After having considered children's proportion of choices as is traditionally done, we will consider how availability interacts with the signal detection theory indexes: sensitivity and response bias. We suggest that multiple comparisons between learning and generalization items help highlight relevant dimensions for generalization and therefore improve children's sensitivity (i.e., ability to discriminate taxonomic items among noise). For these same reasons, response bias (i.e., children's willingness to generalize to taxonomic items) may also be favored by simultaneous settings. However, the effect of sequential settings is an open question. In the present study, we also consider age and conceptual distance between items. We suggest that age and a greater conceptual distance improve children's sensitivity to taxonomic choices.

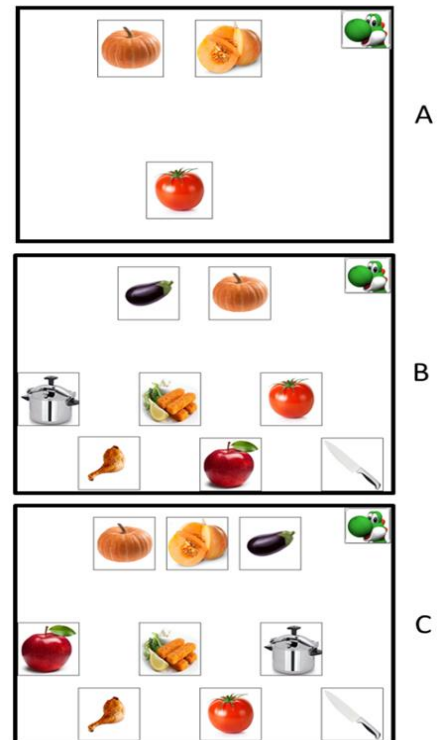


Figure 1: Trial examples built for the food category crossing the three experimental conditions (learning, generalization, availability).

Note: A = close learning - sequential availability, B = far learning - simultaneous availability, C = close-and-far learning - simultaneous availability.

Methods

Participants

Ninety-six French speaking children were tested individually in a quiet room at their school. Two age groups were tested, 47 three-year-old children (mean age: 3.1; range: 2.4 – 3.5) and 49 four-year-old children (mean age, 4.0; range: 3.6 - 4.9). Informed consent was obtained from their school and their parents.

Materials

Color pictures of real objects were used as stimuli. The pictures were organized into fifteen stimulus sets, each associated with a semantic category (e.g., accessories, foods, clothing, tools see Table 1), each set being constructed around learning stimuli and six generalization stimuli. The list of stimuli is given in Table 1.

The design worked as follows. Each participant saw fifteen trials which were divided into three learning conditions (close, far, and close-and-far). Each trial was constructed around a semantic category (e.g., foods, tools, see Table 1). In each learning condition, one of the two pictures was considered as the standard picture. In the close learning condition, the two learning items were two pictures of stimuli from the same basic level category (e.g. a pear and a cut pear). In the far learning condition, the two learning items were pictures of stimuli from the same superordinate category (e.g., a pear and a raspberry). Finally, in the close-and-far learning condition, there were three learning stimuli, two from the same basic level category and one from the same superordinate category (e.g., a pear, a cut pair and a raspberry). As can be seen, the latter condition, was the accretion of the first two conditions. Note that each participant saw the three learning conditions but the semantic categories that composed them differed from one learning condition to the other. In fact, each participant saw only one learning condition from each the 15 semantic categories given in Table 1 (e.g. if he/she saw the pear and the cut pear in the close learning condition, he/she did not see the pear and the raspberry).

Table 1: List of stimuli for each category

	Learning Items			Transfer Items					
	Standard	Close	Far	Near		Distant		Thematically related	
Accessories	Bracelet	Curb-chain	Watch	Pendant	Earring	Bow tie	Hair band	Hand	Girl
Fruits	Pear	Cut pear	Raspberry	Apricots	Pineapple	Chips	Pasta	Basket	Person eating
Tools	Hammer	Hammer 2	Axe	Screw-driver	Pliers	Chain saw	Drill	Building site	Worker
Food	Pumpkin	Pumpkin 2	Eggplant	Tomato	Apple	Chicken	Fish fingers	Knife	Saucepan
Clothing	Sock	Sock 2	Jeans	T-shirt	Jumper	Hat	Tie	Washing machine	Man
Cleaning tools	Broom	Broom 2	Mop	Dustpan	Vacuum cleaner	Washing machine	Dish washer	Cleaning lady	Floor
Fruits 2	Apple	Apple 2	Pineapple	Banana	Orange	Fish fingers	Meat	Knife	Fruit bole
Animals	Snake	Snake 2	Lizard	Turtle	Chameleon	Rabbit	Cow	Aquarium	Land-scape
Food 2	Potato	Potato 2	Salad	Cucumber	Carotte	Piece of tart	Sweets	Bole	Fork
Food 3	Ice cream	Ice cream 2	Cake	Yogurt	Fruit salad	Egg	Sausage	Fridge	Baby eating
Insects	Ladybird	Ladybird 2	Beetle	Butterfly	Ant	Goldfish	Duck	Garden	Sun-flower
Transport means	Bicycle	Bicycle 2	Skateboard	Scouter	Roller blades	Boat	Airplane	Helmet	Bicycle path
Stationary	Pencil	Pen	Ruler	Rubber	Scissors	Computer	Calculator	Classroom	Student
Musical instruments	Guitar	Guitar 2	Drum	Flute	Trumpet	Piano	Stereo machine	Dancer	Music stand
Cuter	Knife	Knife 2	Chopper	Saw	Pincers	Lawn mower	Chain saw	Water-melon	Fork

There were also six generalization items which were two pictures of stimuli from the same superordinate level category as the learning-items (near superordinate generalization items, e.g., apricots and pineapple, in the pear and/or raspberry case), two pictures from a more remote superordinate category as the learning-items (distant superordinate generalization items, e.g., chips and pasta, i.e., from the superordinate category of foods), and two pictures thematically but not taxonomically related with the learning-items (thematically related distractors, e.g., a fruit basket and person eating).

The trials' order during the task was counterbalanced, as was the order in which were presented the different learning conditions. All 15 trials in a task were presented with the same availability. Half of the participants saw the trials in the sequential availability and half in the simultaneous availability. In the sequential availability, the generalization -items' order was randomized between trials. In the simultaneous availability, the order in which each generalization item appeared was randomized. Figure 1 shows 3 examples of trials built using the stimuli from the food category. The pictures were displayed on a 13inch touchscreen laptop. A picture of Yoshi appeared for six seconds on the right-hand side of the screen after all the stimuli pictures had appeared.

We forged 15 different bisyllabic labels (pseudo-words) which are, as shown by Gathercole and

Baddeley (1993), easier to remember than monosyllabic pseudo-words (e.g., buxi, dajo, zatu, xanto, vira). Syllables were of the CV type which is the dominant word structure in French (from Lexique.org, New, Pallier, Brysbaert, & Ferrand, 2004).

Perceptual similarity and semantic similarity ratings were obtained from eighty university undergraduate students (forty students evaluated the perceived similarity and forty others the semantic proximity). For each of the fifteen categories, we assessed the perceptual and semantic similarities between each of the three learning-items and each generalization item, and the three learning items between each other. As expected, close learning items were rated as perceptually more similar and semantically closer to each other than the far learning items to each other ($ps < .05$). The near superordinate generalization items were also rated as perceptually and semantically more similar to the three learning items than the distant superordinate generalization items ($ps < .05$).

Procedure

Participants were seated at a low table, in a quiet room at their school, facing the laptop, next to the experimenter. They were randomly assigned to one of the availability conditions (sequential, or simultaneous). In both conditions, children were shown a soft Yoshi toy and were told a story about

him to introduce the experiment as a game using the following instruction “This is Yoshi, we are going to play with him. But he lives far away from here and speaks a different language. In the game we are going to learn his language.” The experimenter then showed the fifteen trials. In all three learning conditions learning items appeared one by one near superordinate the top of the screen and the experimenter announced their name as they appeared using the instruction: “Yoshi’s mummy says that this is a *buxi*, and this one is also a *buxi*; Yoshi must find other *buxis* for his mummy...”. In the *close-and-far* learning condition the experimenter repeated the second part of the instruction for the close item then for the far item. Then, the generalization items appeared on the lower part of the screen. Depending on the availability they appeared one by one (sequential condition) and the experimenter said “is this a “*buxi*...?” for each of the 6 generalization items, or all together (simultaneous condition) and the experimenter said : : “which ones of these are also *buxis*, show me the *buxis* but not the other things”, In both conditions the experimenter finished the instruction by “...Take your time, don’t make a mistake, don’t give me your answer before Yoshi appears on the screen”. The children were randomly assigned to one of the 6 trial orders, which were seen by the same number of children. The bisyllabic names were assigned randomly to the categories. At the end, the experimenter checked that the child knew the categories, by showing pictures from each trial and asking to name the objects or explain how they were used.

Design

Three- and four-year old children were compared. They were randomly assigned to one of the two availability (sequential, 48 children or simultaneous, 48 children) that was a between subject factor. Age was crossed with availability, learning distance (close, far or close-and-far) and generalization distance (near superordinate or distant superordinate) which were both within-subject factors.

Results

In this study, in a developmental perspective, we assessed whether generalization item’s availability, (simultaneous or sequential), would differentially influence children’s selection of generalization items. It was assumed that sequential generalization would allow less comparisons between stimuli and, thus, would lead to less correct generalizations and that this would interact with semantic distance, in the sense that sequential generalization would affect more negatively the distant superordinate condition.

In the analyses, we kept participants that met the following criteria. Participants had to have chosen at least one item in more than a third of the trials (quantity criterion), and they must not have chosen

all the items in more than two thirds of the trials (selectivity criterion). Under this criterion, four participants in the simultaneous availability were removed from subsequent analyses.

In the analysis, we compared proportions of correct responses to chance (objects = 50%) with t-tests, and used the Bonferroni correction for multiple comparisons (significance set at .0028, divided by the number of comparisons, 18).

In order to test our hypotheses, we ran a four-way ANOVA on the proportion of taxonomic answers, with Age (3 and 4 years) and availability (sequential and simultaneous) as a between-subjects factor, and learning distance (close, far, close-and-far) and generalization distance (near superordinate, distant superordinate) as within conditions.

Since the most important results were interactions, we will report them first. One important result was the interaction between availability and generalization distance, $F(1,88) = 4.72, p < .05, \eta_p^2 = .051$. Figure 2 suggests that this interaction results from a larger difference between the sequential and the simultaneous conditions in the distant superordinate case than in the near superordinate condition. This is compatible with the idea that distant superordinate items were more difficult when they were shown alone.

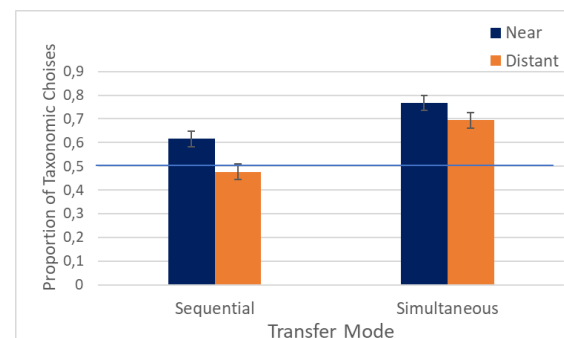


Figure 2: Proportion of taxonomic choices as a function of generalization distance and availability (error bars are SEM).

A posteriori Tukey analysis revealed that in both availability, sequential and simultaneous, participants chose significantly more near superordinate generalization items than distant superordinate generalization items ($p_s < .05$). We also compared performance to chance, with a Bonferroni correction for multiple comparisons (significance at .0028). The t-tests showed that in the simultaneous mode children score above chance in both near superordinate and distant superordinate conditions ($p_s < .001$), whereas in the sequential mode children score above chance in the near superordinate condition ($p < .0025$) but score at chance in the distant superordinate condition ($p = .48$).

Second, availability significantly interacted with age, $F(1,88) = 4.41, p < .05, \eta_p^2 = .048$ (Figure 3). A

posteriori Tukey analysis revealed that the proportion of taxonomic choices in the sequential and the simultaneous modes did not differ in the three-year olds ($p = .61$, $M_{\text{sequential}} = .56$, $M_{\text{simultaneous}} = .65$), but differed significantly at four years of age ($p < .001$, $M_{\text{simultaneous}} = .8$, $M_{\text{sequential}} = .53$). The test also revealed that there was no significant difference between three and four year olds in the sequential condition ($p = .60$) but was marginally significant in the simultaneous condition ($p = .07$). One sample t-tests, with the Bonferroni correction, revealed that all groups were at chance except for 4 year olds in the simultaneous condition who were significantly above chance ($p < .001$). These results show that four year olds benefit from simultaneous generalization but three year olds don't. This is important because it strongly suggests that the comparisons between simultaneous stimuli benefited older children. Whereas the constraints associated with a sequential presentation format were sufficient to impair comparisons between different generalization stimuli.

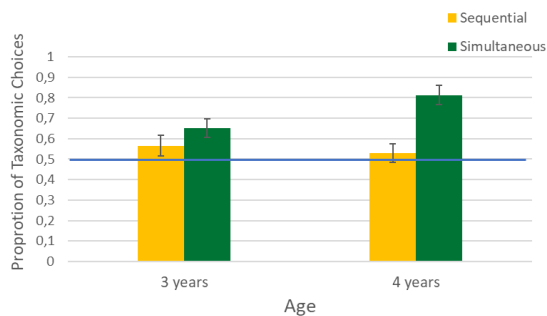


Figure 3: Proportion of taxonomic choices as a function of availability and age (error bars are SEM).

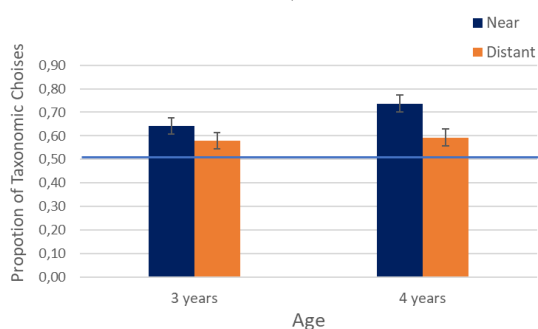


Figure 4: Proportion of taxonomic choices as a function of generalization distance and age (error bars are SEM).

There was an interaction between age and generalization distance, $F(1,88) = 6.28$, $p < .05$, $\eta_p^2 = .067$. A posteriori Tukey test showed that children from each age group, 3- and 4-year olds, chose significantly more near superordinate generalization items than distant superordinate generalization items (respectively $p < .05$ and $p < .001$). Our t-tests revealed that for both age groups, participants score

above chance in the near superordinate condition, ($p < .001$), but participants score at chance in the distant superordinate condition (3-year-olds: $p = .06$; 4-year-olds: $p = .03$).

There was also a learning distance x generalization distance, $F(2,176) = 4.30$, $p < .05$, $\eta_p^2 = .047$, which is interesting but is not central for our study. We also observed main effects, a main effect of availability. There were significantly more taxonomic choices in the simultaneous condition than in the sequential condition, availability, $F(1,88) = 15.16$, $p < .001$, $\eta_p^2 = .15$; $M_{\text{sequential}} = 0.55$, $M_{\text{simultaneous}} = 0.73$. The analysis also showed a significant main effect of learning $F(2,176) = 5.48$, $p < .01$, $\eta_p^2 = .059$; $M_{\text{close}} = 0.61$, $SD_{\text{close}} = 0.28$; $M_{\text{far}} = 0.66$, $SD_{\text{far}} = 0.28$; $M_{\text{close-and-far}} = 0.64$, $SD_{\text{close-and-far}} = 0.29$. A Tukey HSD showed that children gave less taxonomic answers in close learning than in the far learning, $p < .001$ and that close and close-and-far learning did not differ significantly, $p = .15$. Last, we found a significant main effect of generalization distance, $F(2,88) = 49.05$, $p < .001$, $\eta_p^2 = .36$; $M_{\text{near}} = 0.69$, $M_{\text{distant}} = 0.59$.

The difference between the near and the distant superordinate conditions is also confirmed by another dependent that can be derived from our hypotheses. We hypothesized that children in the simultaneous condition, might continue to build their conceptual representation associated with a concept during the generalization phase and that they should start with the near generalization items. We computed the number of near and distant superordinate items in the first two items children selected. A t-test showed that near items were significantly chosen more often than distant superordinate items ($p < .002$, $M_{\text{near}} = 15$, $M_{\text{distant}} = 11$).

We also performed a last control and checked whether the effect of availability might result from a bias to select more items in the simultaneous condition than in the sequential availability. This bias, if any, would result in more thematic lures in the simultaneous condition. An independent samples t-test showed that children gave as many thematically related answers in both availability, ($p > .10$).

Addition to the paper:

The following results include the thematically related item's proportion of choices and the signal detection theory analysis.

Proportion of choices analysis

In order to test our hypotheses, we ran a four-way ANOVA on children's proportion of answers, with age (3 and 4 years) and availability (sequential and simultaneous) as a between-subjects factor, and learning distance (close, far, close-and-far) and generalization item (near superordinate, distant superordinate, thematically related) as within-

subjects factors. Post hoc Tukey comparisons are used to describe all significant effects.

Results revealed simple effects of generalization item $F(2,176) = 169.73, p < .001, \eta_p^2 = .70$ (Figure 5), more near superordinate generalization items were chosen than distant superordinate generalization items ($M_{Near} = 0.69; M_{Distant} = 0.58; p_{Tukey} < .001$) themselves more often chosen than thematically related items ($M_{Theme} = 0.30; p_{Tukey} < .001$).

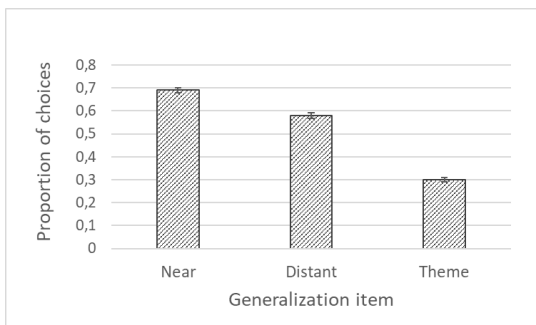


Figure 5: Proportion of choices as a function of generalization item (near superordinate, distant superordinate, thematically related). Error bars are SEM.

Simple effects of learning distance $F(2,176) = 7.22, p < .001, \eta_p^2 = .08$, and availability $F(1,88) = 12.28, p < .001, \eta_p^2 = .12$ were also revealed. We will not describe these effects because they consider taxonomic and thematically related answers together while these items have different statuses and therefore say little of participants taxonomic generalization. However, we can note that results in the close and far learning condition were equivalent to the far learning condition ($M_{Close} = 0.50; M_{Far} = 0.54; M_{C\&F} = 0.53$).

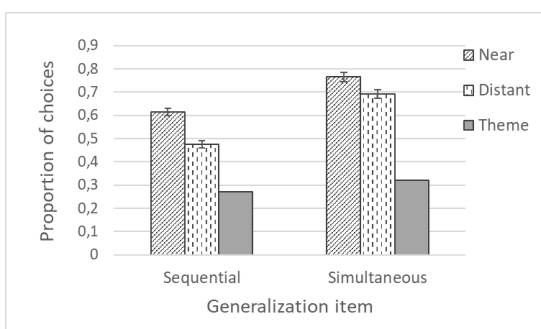


Figure 6: Proportion of choices as a function of generalization item (near superordinate, distant superordinate, thematically related) and availability (sequential, simultaneous). Error bars are SEM.

Results revealed an interaction effect between generalization item and availability $F(2,176) = 8.40, p < .001, \eta_p^2 = .09$, that reveals that children choose more taxonomic items (near superordinate and distant superordinate) in simultaneous availability

settings rather than in sequential availability settings (Figure 6).

Results also revealed interactions between age and availability $F(1,88) = 4.13, p < .05, \eta_p^2 = .05$ that shows that three-year-old children's proportion of choices is not significantly different as a function of availability while four-year-old children choose more items in simultaneous availability settings. We interpret three-year-old children's absence of effect of availability as the sign the cognitive cost of simultaneous settings is too high for their cognitive development. Indeed, simultaneous availability settings likely favor generalization because they allow children to make easier and more comparisons and contrasts between generalization items than sequential settings do. However, the extra comparisons and contrasts must generate cognitive cost that three-year children are less likely to manage than four-year-old children.

Finally, results revealed an interaction between generalization item and age $F(2,176) = 4.68, p < .05, \eta_p^2 = .05$, this interaction however was redundant with the simple effect of generalization item because post hoc tests do not reveal any significant differences between age groups.

From this experiment we will remember that the close and far learning condition was not significantly advantageous compared to the far learning condition. We will not include this learning condition in following studies.

We can remember that from four years of age simultaneous availability of generalization items improves children's proportion of taxonomic choices. This condition appears a favorable setting for taxonomic generalization thanks to the comparisons and contrasts it enables between learning to generalization item mappings.

Signal detection theory analyses

For the data analysis we calculated a sensitivity index D' and a response bias index β derived from signal detection theory (Macmillan & Creelman, 2004) adapting them to experiments based on small numbers of stimuli (see, Rioux et al., 2018b). D' (range 0 : 1) indicates participants ability to discriminate (high values for better discrimination) and β indicates their strategy (liberal, for negative values of β ; conservative for positive values of β ; neutral for $\beta = 0$).

Sensitivity We ran a three-way repeated measure ANOVA on the sensitivity index D' with age (3, 4 years) and availability (sequential, simultaneous) as a between-subject factors, and learning distance (close, far) as a within-subject factor. The simple effect of availability $F(1,92) = 6.09, p < .05, \eta_p^2 = .06$ revealed that children's sensitivity is higher in simultaneous settings ($M_{Simultaneous} = 1.0; M_{Sequential} = 0.74$), and the effect of age $F(1,92) = 6.09, p < .05, \eta_p^2 = .06$ revealed that 4-year-old children's

sensitivity was higher than younger children's ($M_{3yo} = 0.72$; $M_{4yo} = 1.02$).

Bias We ran a three-way repeated measure ANOVA on the bias index Beta with age (3, 4 years) and availability (sequential, simultaneous) as between-subject factors, and learning distance (close, far) as a within-subject factor. The analysis revealed a simple effect of availability $F(1,92) = 14.55$, $p < .001$, $\eta^2_p = .14$ where the bias is conservative in sequential settings and neutral in simultaneous settings ($M_{Simultaneous} = 0.03$; $M_{Sequential} = 0.40$). The analysis also revealed an interaction between availability and age $F(1,92) = 5.50$, $p < .05$, $\eta^2_p = .06$ (Figure 7) where 3-year-old children's bias is the same across availability conditions whereas 4-year-old children's bias is high (more conservative) in sequential than simultaneous settings.

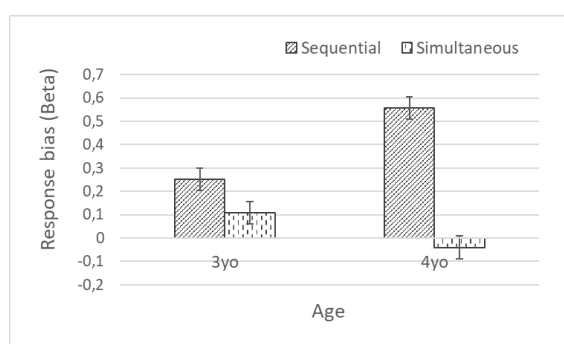


Figure 7: Response bias (Beta) as a function of age (3-, 4-years) and availability (sequential, simultaneous). Errors bars are SEM.

Discussion

Our main question was whether generalization item's availability, sequential or simultaneous, would influence children's selection of generalization items. Related factors such as semantic distance and age were also considered.

Addition to the paper:

Before discussing interactions between availability and other factors it is important to highlight that over all taxonomically based generalization performances were improved in simultaneous availability settings over sequential availability settings. Indeed, proportion of taxonomic answers were systematically higher in simultaneous settings compared to sequential settings.

Our results suggest that availability had an effect and interacted with semantic distance and age. The interaction with generalization distance shows that the difference between *near superordinate* item scores and *distant superordinate* items scores was smaller in the *simultaneous* case. This can be interpreted in terms of comparisons between items

during the generalization view, the distant superordinate items benefiting from the former encoding of the near superordinate items.

The interaction between age and availability, showed that, overall, the *simultaneous* availability favored taxonomic answers and that the difference with the sequential mode was larger for four-year olds. Interestingly, this shows that older children benefited the most of simultaneity which means that they could compare the learning and the generalization stimuli and make sense of all these comparisons. In contrast, the sequential mode remained difficult for them, which means that the comparisons between learning items alone did not lead to better results than in younger children. For the younger children, it might be possible that both situations remained difficult and that they had difficulties also with the simultaneous case and its multiple comparisons between all the available items. Overall, these results show that *simultaneous* availability gives better results and is consistent with the hypothesis that participants benefited from the comparison between *all* the stimuli. These comparisons might contribute to circumscribe the category scope.

The interaction between age and generalization distance showed that children chose less *distant* superordinate generalization items than *near* superordinate generalization items, but the difference between near and distant was larger in the older children. The distance effect is most often interpreted by saying that children find it harder to generalize to a remote item in a semantic taxonomy, or to apply the conceptual criteria to distant superordinate items. The interaction shows that our older children improved first for the near superordinate items. However, these results have a different meaning in a free-choice task. In forced-choice designs, it means that participants, given the learning input they receive, do not understand the targeted concept. In a free-choice design, this means that participants do not consider that an item belongs to the category. They might have forged a conceptual representation that does not include these items.

Another interesting result regarding semantic distance is the interaction between learning distance and generalization distance. There was a linear trend in the near superordinate generalization stimuli from close to close-and-far, whereas performance in the *distant* superordinate generalization declined in the close-and-far learning case, resulting in a larger difference between near superordinate and distant superordinate generalization stimuli in the latter case. This suggests that integrating a *close* and a *far* learning item during comparison might have elicited integration difficulties in children, that is difficulties to consider that a *close* and a *far* learning item belong to the same class (e.g. an apple and an apple belong to the same basic level category, whereas an apple and a banana belong to the same superordinate

category). Integrating two conceptual distance levels in a single representation might be challenging.

Addition to the paper:

Results from signal detection theory analyses confirmed the positive effect of simultaneous generalization item availability on taxonomic generalization. First, children's sensitivity to taxonomic items was improved in simultaneous settings over sequential settings. This result supports the suggestion that generalization items influence the category representation used during generalization; It can be suggested that improved sensitivity is due to a denser representation that may have been built in part thanks to generalization item information.

These analyses also reveal that children's response bias is neutral in simultaneous settings. Children therefore have neither reluctance nor over confidence when generalizing to taxonomically related items in simultaneous setting whereas they show reluctance to do so in sequential settings were their response bias is conservative. A possible interpretation of this result is that the comparisons that can be made between generalization support generalization according to a taxonomic categorization rule. How this may be obtained however is an open question.

One interpretation of these results is that comparisons between learning and generalization items help children to generalize because it enables them to build a deeper representation of the targeted relation between the learning items. Thus, the test phase also contributes to the extraction of the targeted concept and the building of the concept's representation. Children might progressively align the learning items with the generalization items (Gentner & Colhoun, 2010). These learning-item to generalization -item comparisons, starting with near superordinate generalization items, may progressively enable children to extend the scope of their representation and choose more distant superordinate items.

This interpretation implies that children may use information from multiple sources while learning and generalizing novel words. However, integrating multiple information depends on proficient executive functions, functions that are not fully developed around 3 to 4 years age at which children build representation to generalize novel words. Complex interactions between amount of information available to fulfill the task and the tasks executive difficulty may exist as Augier & Thibaut (2013) study show with their manipulation of multiple item comparisons. Future studies could focus on these complex interactions that may exist between representation building rich information environments and executive function limits.

Acknowledgment

We wish to thank the french Agence Nationale pour la Recherche (ANR) for their financial support in the COMPARE project to the third author. The first author benefited from a French ministry for research scholarship. We also wish to thank Flavie Champaney and Virginie Pickel for their help in the stimulus construction and data collection.

References

- Augier, L., & Thibaut, J.-P. (2013). The benefits and costs of comparisons in a novel object categorization task: Interactions with development. *Psychonomic Bulletin & Review*, 20(6), 1126-1132. <https://doi.org/10.3758/s13423-013-0436-x>
- Childers, J. B., & Paik, J. H. (2009). Korean- and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of child language*, 36(1), 201-224. <https://doi.org/10.1017/S0305000908008891>
- Christie, S., & Gentner, D. (2010). Where Hypotheses Come From: Learning New Relations by Structural Alignment. *Journal of Cognition and Development*, 11(3), 356-373. <https://doi.org/10.1080/15248371003700015>
- Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Dev*, 1173-1188.
- Gentner, D., & Colhoun, J. (2010). Analogical Processes in Human Thinking and Learning. In B. Glatzeder, V. Goel, & A. Müller (Éd.), *Towards a Theory of Thinking* (p. 35-48). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03129-8_3
- Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison Facilitates Children's Learning of Names for Parts. *Journal of Cognition and Development*, 8(3), 285-307. <https://doi.org/10.1080/15248370701446434>
- Gentner, D., & Namy, L. L. (1999). Comparison in the Development of Categories. *Cognitive Development*, 14(4), 487-513. [https://doi.org/10.1016/S0885-2014\(99\)00016-7](https://doi.org/10.1016/S0885-2014(99)00016-7)
- Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference? *Cognition*, 112(1), 105-119. <https://doi.org/10.1016/j.cognition.2009.03.012>
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Smiley, S. S., & Brown, A. L. (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology*, 28(2), 249-257.
- Thibaut, J.-P., & Witt, A. (2015). Young children's learning of relational categories: Multiple

- comparisons and their cognitive constraints. *Frontiers in psychology*, 6, 643.
- Thibaut, J.-P., & Witt, A. (2017). Generalizing novel names in comparison settings : Role of conceptual distance during learning and at test. Proceedings of the 39th Annual Meeting of the Cognitive Science Society (pp. 3314-3319).
- Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, 36(5), 571-581.

Chapter 2 - How stimuli availability effects novel noun generalization in a free-choice design.

The following chapter contains the proceedings paper: Stansbury E., Witt A., Thibaut J-P., (2022). How Stimuli Availability Effects Novel Noun Generalization in a Free-Choice Design. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 44th, pp. (2789-2795).

Note: Amendments made to the paper for the clarity of the thesis are in boxes.

How Stimuli Availability Effects Novel Noun Generalization in a Free-Choice Design.

Eleanor Stansbury (eleanor.stansbury@u-bourgogne.fr)

Arnaud Witt (arnaud.witt@u-bourgogne.fr)

Jean-Pierre Thibaut (jean-pierre.thibaut@u-bourgogne.fr)

LEAD-CNRS, UMR 5022, Université de Bourgogne Franche-Comté,
Pôle AAFE – Esplanade Erasme, 21065 Dijon, France

Abstract

A common result in novel word generalization is that comparison settings (i.e., several stimuli introduced simultaneously) favor conceptualization and generalization. We investigated which type of items four-, five- and six-year-old children would choose as referents in a free-choice novel noun generalization task. We manipulated the generalization items availability at test (i.e., generalization stimuli introduced sequentially or simultaneously). We also manipulated the semantic distance between items. In a signal detection theory framework, results showed that a simultaneous presentation of generalization items improves children's sensitivity and helps them use a neutral strategy to generalize. Conceptual distance at learning also affects generalization performance. We discuss the cognitive constraints that both types of presentation bring into the task, and how distance might impede or favor conceptual alignments.

Keywords: Categorization, generalization, novel noun, forced-choice, free-choice, conceptual distance, stimuli availability

Introduction

When children learn to categorize and name novel objects, they have to understand which dimensions are important to define the corresponding concept (Murphy, 2002). Identifying which word learning format(s) promote concept construction and novel word generalization is an important topic for cognitive sciences.

The present study capitalizes on recent evidence showing that comparing stimuli from the same category during learning favors conceptually based novel word generalization. However, the benefits of comparison have mainly been evidenced with forced-choice design. In contrast, we used a free-choice task and manipulated the “temporal” availability of the generalization stimuli, either *sequential* (one-by-one) or *simultaneous* (all stimuli displayed together), as a function of semantic distance. We analyzed the answers, in a signal detection theory (SDT) framework, in order to study sensitivity and response bias across ages and the generalization stimuli's availability.

Addition to the paper:

Analyses of children's generalization performances have also been added to the results.

Comparisons and novel word generalization

Recent evidence shows that the opportunity to compare exemplars while learning a novel word favors conceptually

based categorization and the novel word generalization compared to the classical single exemplar learning design. This result has been replicated in various linguistic categories (see Gentner & Christie, 2010 for a review).

In a typical comparison choice design, the learning items are both perceptually similar (i.e. they display the same shape, e.g. two similar fruits) and taxonomically similar. In the generalization phase, the child has to choose between an item that is taxonomically-related to the learning items but rated as perceptually dissimilar to them (e.g., a banana) and a perceptually similar lure that is conceptually unrelated to the learning items but perceptually similar (e.g., a red Christmas ball) to them (Gentner & Namy, 1999). Whatever the variations and particulars of the design, these studies have shown that comparison situations and the presence of a unifying name (i.e. label effect) lead to more conceptually-based generalizations than no-comparison situations, which, in turn, tend to favor the perceptually salient (the shape) lure.

The semantic distance between the items compared has also been shown to have an important effect on children's taxonomic generalizations. Thibaut and Witt (2017) studied novel noun generalization with 4- and 6- year-old children. They manipulated the semantic distance between learning items (close vs. far) and between the learning items and the taxonomic item to generalize the word to (near superordinate vs. distant superordinate). Close learning items were from the same basic level category whereas far learning items were from the same superordinate category. Results revealed that learning pairs from more distant domains led to better taxonomic generalization. The interpretation was that broader conceptual distance at learning helps participants abstract the relevant relations between learning objects and build a more conceptually based representation for generalizing the novel word.

Similar effects of distance between learning items that favor generalization have been demonstrated in various domains like relational noun generalization (Thibaut & Witt, 2015), analogical reasoning (Thibaut, French, Vezneva, 2010), multidimensional stimuli categorization (Hammer et al. 2008), or scientific reasoning (Klahr & Chen, 2011).

Forced-choice and free-choice tasks

Most existing studies on object noun comparison and generalization use a forced-choice design (Alfieri et al., 2013). Forced-choice designs are well-suited to study

children's biases in word learning tasks (Landau et al., 1988; Markman, 1989) or to study which commonalities children spontaneously choose as a basis for generalization.

However, in such designs children might also choose the item that is most plausibly related to the learning item(s) but that they would not select as an item of the same category if they were not forced to choose. Conversely, selecting one option does not mean that participants would not accept the other option as a member of the category. For example, in a forced-choice, Smiley and Brown (1979) showed that young children could select and justify a taxonomic choice even when their first choice was a thematic choice.

Free-choice designs can therefore be considered better suited to study the extension children give to novel words because children can select all or none of the generalization options, whatever their conceptual relation with the training items (perceptual lure, taxonomically related choice, or theme related choice). Therefore, free choice designs might give a clearer picture of the items children believe belong to the category.

Goals of the present experiment

We use a free-choice task with 4- to 6-year-old children with the following questions. First, will comparison, semantic distance and attractive lures still affect children's conceptually based generalization in a free choice design? Indeed most of the available evidence has been obtained under "forced" circumstances? Indeed, it might be argued that only close items will be accepted as "natural" extension of terms when given the possibility to overlook items. Second, how will the generalization item's availability - a new parameter that can only be considered in free-choice designs and the effect of which has not yet been analyzed - influence generalization and interact with parameters already recognized as important during generalization?

Indeed, in daily life, children might see generalization items simultaneously (e.g., a fruit with other target fruits, or a fruit in a kitchen with other fruit-related objects, or a fruit with a perceptually similar object that is not a fruit). They might also encounter the same generalization items one by one, sequentially. We therefore used two different generalization availability conditions (simultaneous and sequential) to present the generalization items. These items were selected to approximate the diversity of generalization items children may encounter in daily life. They were taxonomically related, thematically related, perceptually related or non-related to the learning items.

Addition to the paper:

First, we analyzed children's proportion of choice profiles as a function of the tasks factors (age, availability, learning distance).

Because we used a free-choice design, we analyzed our results in the signal detection theory framework, analyzing children's sensitivity to the signal, (i.e., ability to discriminate

between taxonomically related items, the signal, and the other generalization items, the noise) and children's response strategy (i.e., children's willingness or confidence in the chosen generalization rule). Sensitivity is distinguished from their response strategy known as the response bias, which can be conservative (i.e., tendency to reject answers that leads to missing correct answers), neutral or liberal (i.e., tendency to accept, leading to many false alarms). Combining different measures of performance (e.g., correct answers, false alarms) in one index will give us information that percentage of choices do not provide. Indeed, the number of incorrect choices will be taken into account by the inclusion of false alarms. It is also interesting to consider the extent to which different conditions might or might not boost the overall probability of accepting.

In this experiment we also manipulated age as a between factor, predicting that sensitivity should increase and bias decrease with age; and conceptual distance between the compared items in order to study whether this might interact with availability and influence bias and sensitivity.

We predicted that a larger semantic distance between learning items would lead children to build a broader learning representation. In the far learning condition children should therefore include more distant superordinate generalization items and have a higher level of sensitivity.

We also predicted that the generalization item's availability will influence children's sensitivity because the two availability conditions (i.e., sequential or simultaneous) constrain the task - and children's generalization - differently. However, how generalization item availability will affect sensitivity is an open question. The two availability conditions enable different comparisons between learning and generalization items that children may use to find their answer. On the one hand, sensitivity may be lower in the simultaneous condition, because children's attention may be attracted by the perceptual items and diverted from the taxonomic answers reducing sensitivity, whereas no such interference is possible in the sequential condition. On the other hand, sensitivity might be higher in the simultaneous condition, because multiple comparisons between learning and generalization items might help to highlight conceptual commonalities and reject irrelevant dimensions, thus reducing false alarms.

As for children's strategy (response bias), we predict that the simultaneous condition should reduce the bias compared to the sequential one because of the possibility to compare all the stimuli and decide which items belong to the category in terms of a reference set of features.

We will also follow up with a control no-comparison situation in a second step. Our aim is to have results in a no-comparison design as reference in a free choice design, because this type of data is not available in the literature.

Experiment 1

Methods

Participants

One-hundred-and-eighty French speaking children were tested individually in a quiet room at their school. Informed consent was obtained from their parents. Three age groups were tested, 58 four-year olds (47 months; 41 – 53), 62 five-year-olds (60 months; 56 – 65) and 60 six-year-olds (72 months; 68 – 83).

Materials

Color pictures of real objects were used as stimuli. The pictures were organized into sixteen stimulus sets, each associated with a semantic category (e.g., accessories, foods, clothing, tools, etc), each set was designed with three learning stimuli and ten generalization stimuli. The sixteen trials were divided into two learning conditions (close or far learning). Each trial was constructed around a semantic category. In each learning condition, one of the two pictures was considered as the standard picture. In the close learning condition, the two learning items were two pictures of objects from the same basic level category (e.g. a pear and a cut pear). In the far learning condition, the two learning items were from the same superordinate category (e.g., a pear and a raspberry).

The ten generalization items were : two pictures of objects from the same superordinate level category as the learning-items (near superordinate generalization items, TaN, e.g., apricots and pineapple), two pictures from a more distant superordinate category as the learning-items (distant superordinate generalization items, TaD, e.g., chips and pasta); two stimuli perceptually similar to the standard learning item but not taxonomically related to the learning items (perceptual distractors, P, e.g. a punching ball and a pear shaped candle); two pictures thematically but not taxonomically related to the learning-items (thematically related distractors, Th, e.g., a fruit basket and a fruit knife); two lures semantically and perceptually unrelated to the learning items (non-related distractors, NR, e.g. a car and a note book). Twenty 3-year-old and twenty 4-year-old children were asked to recognize the trials objects and succeeded, controlling that all items are known by children.

The trials' order during the task was balanced, as was the order in which were presented the different learning conditions. All 16 trials in a task were presented with the same generalization availability. In the sequential generalization availability, the generalization-items' order was balanced between trials. In the simultaneous generalization availability, the position of the generalization items on the screen appeared was balanced between trials. Generalization availability was set as a between factor to avoid that answers given in simultaneous cases influence answers given in sequential cases, if for example taxonomic answers are more obvious in simultaneous cases were all items are available together. Figure 1 shows an example of a trial built using the stimuli from the food/fruit/pear category. The pictures were displayed on a 13inch touchscreen laptop.

We forged 16 different bisyllabic labels (pseudo-words) which are, as shown by Gathercole and Baddeley (1993), easier to remember than monosyllabic pseudo-words (e.g.,

buxi, dajo, zatu, xanto, vira). Syllables were of the CV type which is the dominant word structure in French (from Lexique.org, New, Pallier, Brysbaert, & Ferrand, 2004).

Ratings on a 1 to 10 scale (1: far rating, 10: close rating) were obtained from undergraduate students to control generalization items. Twenty-eight students' ratings confirmed that taxonomically related items are considered to belong to the same category as the standard learning item (average ratings: Ta: 7.6, Th:4.5, P:2.1, NR:1.5, average p between Ta-Th $p < .001$). And twenty-four students rated near superordinate taxonomically related generalization items conceptually

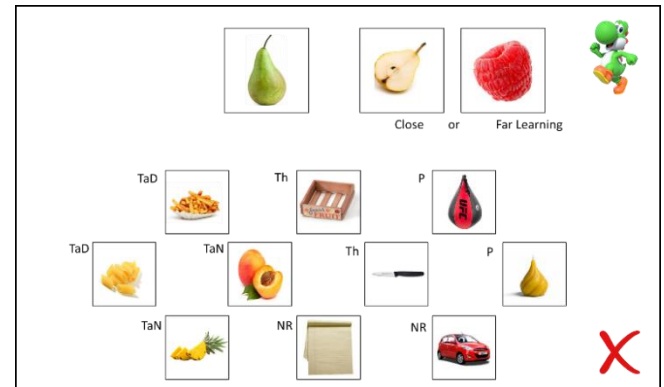


Figure 1: Trial built for the food category

Note: Participant saw either the close or the far learning item
TaN : near superordinate, TaD: distant superordinate, Th: thematically related, P: perceptually related, NR : non related generalization items

closer to the standard item compared to distant taxonomically related items.

Perceptual similarity and thematic similarity ratings from 36 and 21 students respectively, controlled that the item were perceptually more similar (average ratings: Ta: 3.0, Th:2.2, P:6.3, NR:1.7, average p between P-Ta $p < .001$) or thematically more strongly related (average ratings: Ta: 6.4, Th:7.5, P:2.2, NR:1.8, average p between Th-Ta $p < .05$) to the standard learning item than the taxonomically related items. Unrelated distractors scored significantly below all other generalization items in all ratings ($p < .01$)

Procedure

Participants were seated at a low table, in a quiet room at their school, facing the laptop, next to the experimenter. They were randomly assigned to one of the generalization availability conditions (sequential, or simultaneous). In both conditions, children were introduced to a puppet named “This is Yoshi, we are going to play with him. But he lives far away from here and speaks a different language. In the game we are going to learn his language.” The experimenter then showed the fifteen trials. In all two learning conditions learning items appeared one by one near the top of the screen and the experimenter announced their name as they appeared using the instruction: “Yoshi’s mummy says that this is a *buxi*, and

this one is also a *buxi*; Yoshi must find other *buxis* for his mummy....”. Then, the generalization items appeared on the lower part of the screen, generalization availability one by one in the sequential condition, the experimenter said “is this a “*buxi*...?” for each of the 10 generalization items. In the simultaneous condition, they were displayed simultaneously:

“which ones of these are also *buxis*, show me the *buxis* but not the other things”. The experimenter finished the instructions by “Take your time, don’t give me your answer before Yoshi appears on the screen”.

Table 1: Mean proportion of answers as a function of generalization items, generalization availability and age

Availability	Sequential						Simultaneous					
	Close			Far			Close			Far		
Learning distance												
Age	4	5	6	4	5	6	4	5	6	4	5	6
Near	.44 (.04)	.49 (.04)	.60 (.05)	.44 (.04)	.57 (.04)	.63 (.05)	.64 (.04)	.75 (.04)	.72 (.05)	.65 (.04)	.75 (.04)	.72 (.05)
Distant	.44 (.03)	.37 (.04)	.45 (.05)	.48 (.03)	.34 (.04)	.46 (.05)	.55 (.03)	.51 (.04)	.52 (.05)	.58 (.03)	.57 (.04)	.51 (.05)
Thematically related	.38 (.03)	.24 (.03)	.33 (.04)	.26 (.03)	.10 (.03)	.19 (.04)	.30 (.03)	.27 (.03)	.17 (.04)	.32 (.03)	.22 (.03)	.17 (.04)
Perceptually related	.33 (.03)	.22 (.03)	.26 (.04)	.28 (.03)	.19 (.03)	.15 (.04)	.44 (.03)	.35 (.03)	.26 (.04)	.38 (.03)	.29 (.03)	.30 (.04)
Non-related	.27 (.03)	.07 (.03)	.11 (.05)	.19 (.03)	.07 (.03)	.06 (.05)	.38 (.03)	.23 (.03)	.15 (.05)	.37 (.03)	.22 (.03)	.15 (.05)

Note : Means and *Standard deviations in brackets.*

Generalization item (Near superordinate, Distant superordinate, Thematically related, Perceptually related, Non-related) Generalization availability (Sequential or Simultaneous), Learning distance (close, far), Age (4-, 5-, or 6-years-old).

Table 2 : Proportion of responses and signal detection indexes as a function of the comparison situation, learning distance and generalization availability

Comparison situations	Learning distance	Generalization availability	Hits	Misses	False Alarms	Correct rejections	D'	β
			Comparison situations	Close	Sequential	.43	.57	.17
Simultaneous	.61	.39			.27	.73	.83	.23
Far	Sequential	.45		.55	.24	.76	.47	.42
	Simultaneous	.63		.37	.28	.72	.85	.18
No-Comparison situations	NA	Sequential	.17	.83	.06	.68	.23	1.63
		Simultaneous	.49	.51	.32	.94	.35	.40

Design

Four-, five- and six-year-old children were compared. They were randomly assigned to one of the two generalization availabilities (sequential, 88 children or simultaneous, 92 children) a between subject factor. Age was crossed with generalization availability, and learning distance (close, far) a within-subject factor.

Results

Proportion of choice analysis

Addition to the paper:

We ran a four-way ANOVA on children's proportion of answers, with age (3 and 4 years) and availability (sequential and simultaneous) as between-subjects factors, and learning distance (close, far) and generalization item (near taxonomic, distant superordinate, thematically related, perceptually related, unrelated) as within-subjects factors. Post hoc Tukey comparisons were used to describe significant effects.

The analysis revealed a simple effect of generalization item $F(4,696) = 257.65, p < .001, \eta_p^2 = .60$, revealing that the differences between the proportion of answers of each type of generalization item were significantly different to the other items' proportion of answers ($p_{\text{Tukey}} < .001$) (Figure 2).

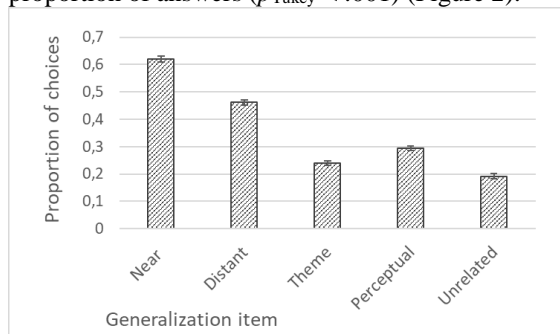


Figure 2: Proportion of answers as a function of generalization item (Near superordinate, distant superordinate, thematically related, perceptually related, unrelated). Error bars are SEM.

Other simple effects were revealed, one of learning distance $F(1,174) = 7.22, p < .01, \eta_p^2 = .04$ revealing that children make more choices in close learning settings rather than far learning settings ($M_{\text{Close}} = 0.37$; $M_{\text{Far}} = 0.35$); and another of availability $F(1,174) = 11.91, p < .001, \eta_p^2 = .06$ that reveals that children choose more items in the simultaneous settings than in the sequential settings ($M_{\text{Simultaneous}} = 0.42$; $M_{\text{Sequential}} = 0.31$). These results however have little meaning because they consider taxonomic items and distractors together while these items have different statuses.

The analysis revealed an interaction effects between learning distance and availability $F(1,174) = 9.22, p < .01, \eta_p^2 = .05$ (Figure 3) showing that the children's proportion of answers were higher in simultaneous settings than in sequential settings ($p_{\text{Tukey}} < .05$), and more importantly that their proportion of answers were different across learning distance conditions but only in sequential settings (Simultaneous : $M_{\text{Close}} = 0.41$; $M_{\text{Far}} = 0.42$; Sequential : $M_{\text{Close}} = 0.33$; $M_{\text{Far}} = 0.29$; $p_{\text{Tukey}} < .001$). In other words, children's generalization performances were affected by learning distance in sequential settings and a greater distance at learning reduced children's proportions of taxonomic generalizations.

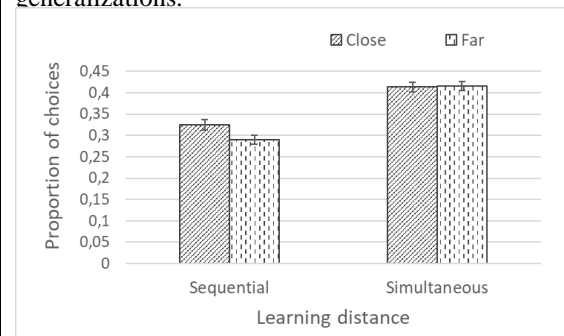


Figure 3: Proportion of answers as a function of learning distance (close, far) and availability (simultaneous, sequential). Error bars are SEM

Other interaction effects were revealed. The interaction between generalization item and availability $F(4,696) = 8.24, p < .001, \eta_p^2 = .05$ revealed higher proportion of choices of taxonomic items in simultaneous settings compared to sequential settings ($p_{\text{Tukey}} < .01$) but equivalent proportion of answers for distractors. This is interesting because it reveals that simultaneous settings do not lead only to an increase in answers for all generalization items but specifically favor taxonomic generalization.

Results also revealed an interaction between generalization item and age $F(8,696) = 10.51, p < .001, \eta_p^2 = .12$ and between generalization and learning distance $F(4,696) = 11.44, p < .001, \eta_p^2 = .06$ however these interactions were redundant with the simple effect of generalization item because post hoc tests did not reveal any significant differences between age groups or between learning distant conditions.

Finally a three way interaction between learning distance, generalization item and availability $F(4,696) = 8.90, p < .001, \eta_p^2 = .05$ was significant. The differences between availability conditions and learning distance conditions were minor.

To summarize, these analyses reveal a well-executed task (see generalization item effect), and two important effects/ One shows that learning distance affects children generalization behavior in simultaneous settings only. The other confirms the

positive effect of simultaneous generalization item availability on taxonomic generalization performances.

Signal detection theory indexes

For the data analysis we calculated a sensitivity index D' and a response bias index β derived from signal detection theory (Macmillan & Creelman, 2004) adapting them to experiments based on small numbers of stimuli (see, Rioux et al., 2018b). D' (range 0 : 1) indicates participants ability to discriminate (high values for better discrimination) and β (range -1 : 1) indicates their strategy (liberal, for negative values of β ; conservative for positive values ; neutral for $\beta=0$).

Data analysis

Proportion of choices for each type of generalization item are given as an indication in Table 1.

Sensitivity. In order to test our hypothesis on the learning distance and availability effects on children's sensitivity we ran a three-way repeated measure ANOVA on the sensitivity index D' with age (4, 5, and 6 years), generalization availability (sequential and simultaneous) as between factors and learning distance (close and far) as a within factor (see Table 2). Results revealed simple effects of age $F(2,174) = 11.99, p < .001, \eta_p^2 = .11$; and availability $F(1,174) = 12.98, p < .001, \eta_p^2 = .07$ and learning $F(1,174) = 24.27, p < .001, \eta_p^2 = .11$. Children's sensitivity increased with age ($M_4 = .47, M_5 = .83; M_6 = .87$). A posteriori Tukey analysis revealed that sensitivity at 4 years is significantly lower than sensitivity at 5 and 6 years ($p_{Tukey} < .001$). Sensitivity was also significantly higher in simultaneous settings compared to sequential settings ($M_{Simultaneous} = 0.86; M_{Sequential} = 0.59$) and after far learning compared to close learning ($M_{Close} = 0.63; M_{Far} = 0.82$).

The analysis also revealed an interaction effect between learning distance and availability: $F(1,174) = 7.60, p < .01, \eta_p^2 = .04$ (Figure 4). A posteriori Tukey comparisons revealed that children's sensitivity (i.e. ability to discriminate the taxonomic items among the distractors) in sequential settings is significantly higher after far learning compared to close learning (Sequential: $M_{Far} = .73, M_{Close} = .44, p_{Tukey} < .05$). In simultaneous settings sensitivity doesn't vary as a function of learning distance.

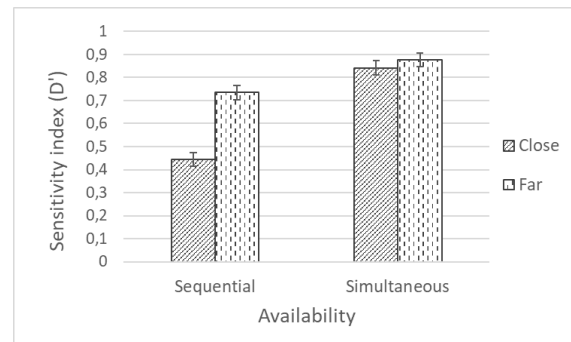


Figure 4 : Sensitivity as a function of learning distance (close, far) and availability (sequential, simultaneous). (error bars are SEM)

Response bias. In order to test the factors' effects on children's strategy, we ran a three-way repeated measure ANOVA on the Bias criterion Beta with age (4, 5, and 6 years) and availability (sequential and simultaneous) as between-subject factors and learning distance (close and far) as a within factor. Results revealed a simple effect of availability $F(1,174) = 14.02, p < .001, \eta_p^2 = .08$ where response bias is more conservative in sequential settings than in simultaneous settings ($M_{Simultaneous} = 0.18; M_{Sequential} = 0.54$) (Figure 5).

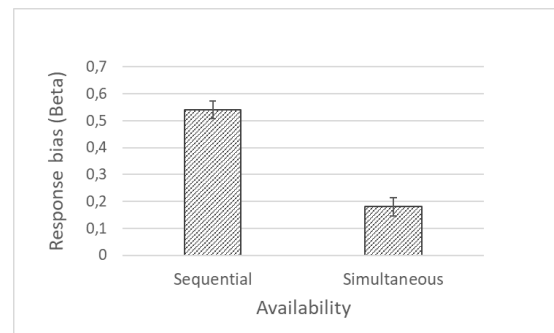


Figure 5: Response bias BETA as a function of availability (sequential, simultaneous).

At this stage it is important to summarize what sensitivity and response bias analyses say of the effects of availability and learning distance on novel noun generalization. Simultaneous settings favor both children's sensitivity and bias, improving sensitivity and leading to a neutral bias.

The effect of far learning distance is of particular interest. In far learning settings compared to close learning settings children's sensitivity increases (i.e. making taxonomic generalization more difficult) and their response bias is less conservative (i.e., making taxonomic generalization more likely). We suggest that the representation built from far learning items is built with more conceptually based items or that the conceptually based items are more weighted than after learning from close items where perceptual properties are favored. The representation being conceptually stronger, sensitivity to other conceptually related items is improved. Bias may be

improved because it is clearer that the categorization rule to follow is a taxonomic rule.

Control Experiment

Experiment 1 examined how children might generalize novel nouns in free choice designs. It aimed to analyze the effect of a parameter already known to affect conceptually based generalization in forced choice designs – learning distance - and the effect of a new parameter peculiar to free choice designs – generalization item availability.

However, experiment 1 tells nothing about the distribution of children's choices in a no-comparison and free-choice design. However, no-comparison situations are an important starting point in most novel noun learning paradigms. Moreover, this study aims to give, a comprehensive description of conceptually based generalization in free-choice designs and, for this reason, knowing how children generalize in a no-comparison situations is an important reference to have.

The following control experiment condition addresses this question. We focused on the age of 5 only because there was no interaction involving age in both sensitivity and bias, in the above analyses.

Methods

Participants

Forty-one, children were tested in a no-comparison situation. Children's average age was 5 years (mean: 56 months, range: 48-70). They were randomly assigned to a sequential (20 children) or simultaneous (21 children) availability condition.

Materials and Procedure

Materials and procedure were similar to the main experiment except that there was only one learning item and thus, the learning distance factor disappears.

Design

Five-year-old children in a no-comparison situation were randomly assigned to one of the two generalization availabilities (sequential, 20 children or simultaneous, 21 children) and were compared with the five-year-old children in a comparison situation. Comparison was crossed with generalization availability as a within-subject factor.

Results

We performed the same analyses as in experiment 1.

Addition to the paper:

Proportion of choices analysis

We ran a three-way ANOVA on children's proportion of choices with availability (sequential and simultaneous) and comparison setting (comparison, no-comparison) as between-subject factors, and generalization item (near superordinate, distant superordinate, thematically related, perceptually related, unrelated) as within-subjects factors. The effects of interest are those involving comparison setting. We are mainly interested in differences between scores in comparison versus no-comparison learning settings. To describe the effects we shall therefore use post hoc comparisons with a Bonferroni significance level set at $p = 0.05/5 = .01$.

The analysis revealed a simple effect of comparison setting $F(1,109) = 7.15, p < .01, \eta_p^2 = .06$ and an interaction between generalization item and comparison setting $F(4,436) = 18.60, p < .001, \eta_p^2 = .15$ (Figure 6). In the comparison condition children's proportion of near superordinate choices and distant superordinate choices were significantly higher than in no-comparison settings ($p < .01$). Comparison at learning favors taxonomic generalization in forced-choice designs (Gentner & Namy, 1999), and as expected it is also the case in free-choice designs. All proportion of choices of distractors are equivalent in comparison and no comparison settings.

The analysis also revealed simple effects of generalization item $F(4,436) = 81.93, p < .001, \eta_p^2 = .43$, and availability $F(1,109) = 24.33, p < .001, \eta_p^2 = .18$.

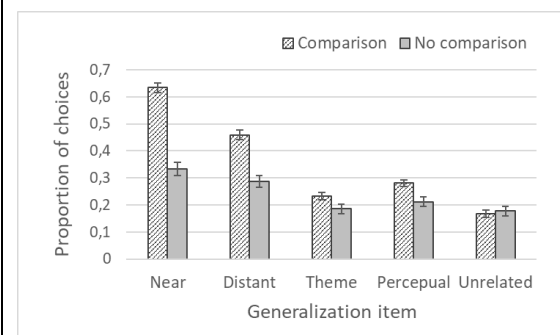


Figure 6: Proportion of choices as a function of generalization item (near superordinate, distant superordinate, thematically related, perceptually related, unrelated) and comparison setting (comparison, no comparison). Error bars are SEM.

Signal detection theory

Sensitivity. The two-way ANOVA on the sensitivity index D' with comparison (comparison, no-comparison) and availability (sequential and simultaneous) revealed an effect of comparison setting $F(1,109) = 16.47, p < .001, \eta_p^2 = .13$.

Children's sensitivity is significantly higher in the comparison situation compared to the no-comparison situation ($M_{comp} = 0.81$ $M_{NoComp} = 0.33$).

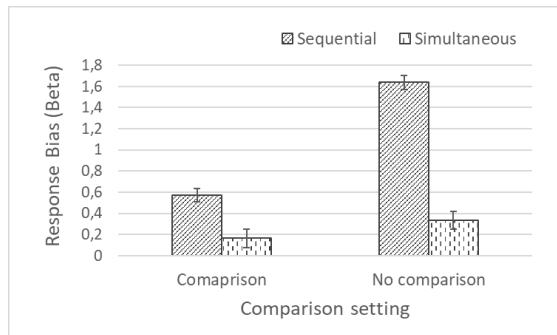


Figure 7: Response bias index Beta as a function of availability (sequential, simultaneous) and comparison setting (comparison, no comparison). Error bars are SEM.

Response bias. A two-way ANOVA on the response bias index β with comparison (comparison and no-comparison) and generalization availability (sequential and simultaneous) as between factors revealed an effect of comparison setting $F(1,109) = 17.00$, $p < .001$, $\eta_p^2 = .14$, of availability $F(1,109) = 32.33$, $p < .001$, $\eta_p^2 = .23$ and an interaction between both factors $F(1,109) = 8.86$, $p < .01$, $\eta_p^2 = .08$ (Figure 7). The interaction and Tukey post hoc comparisons reveal that children response bias is less conservative in no comparison simultaneous cases compared to no comparison sequential settings ($p < .001$).

Discussion

We used a free-choice design in order to study children's word extension rather than a traditional forced-choice word generalization design. These word extensions in a free-choice design are a good indicator of which items children believe belong to the same category as the items the word was learnt with.

Our specific aim was to assess how children would extend a novel noun in this type of design and how their performances would be influenced by the task's factors: conceptual distance between learning items and generalization items' availability.

Addition to the paper:

Analyses of the proportion of taxonomic choices participants made revealed that simultaneous availability of generalization items improved children's proportion of taxonomic answers over sequential availability settings. This result supports the suggestion that generalization items are used during the task and tells us that they help children build taxonomically based representation that are necessary for correct noun generalizations.

A surprising result was the interaction between learning distance and generalization item's availability. It has previously been shown in forced choice designs that children succeeded less well in generalizing to taxonomically related items in far learning settings than in close learning settings (Thibaut & Witt, 2023). Here, this result is confirmed only in sequential availability settings. In simultaneous settings, generalization results are higher, and the learning condition no longer affects performances. Learning distance effects, the representations children build before generalizing the word. If the learning items are conceptually further apart, it is harder for children to build their representation. It seems here, that generalization items available simultaneously help compensate this difficulty. This result would thus argue for the suggestion that generalization items are used for building the representations used during generalization and that they add to the information found in the learning items.

Our main result was that learning distance and generalization availability interacted for both sensitivity and response bias indexes which confirmed our main hypothesis that these factors would affect novel word generalization. For both measures (i.e., sensitivity and response bias) children performed better in the simultaneous condition; their sensitivity was higher, and they were less biased (see Figure 4 and 5).

Performance differences between close and far learning cases only appeared in the sequential condition. In far learning trials, children's sensitivity was higher and their response bias was lower than in close learning trials.

In a control experiment we analyzed results from a no-comparison noun generalization situation versus results from our comparison situations: children were less sensitive as they rejected 71% of taxonomically related items. They were also more biased (i.e., extremely conservative strategy) in the no-comparison situation. Thus, in the no-comparison condition, they took fewer items, mostly incorrect perceptual lures.

Informal remarks at the end of the no-comparison condition suggest that children would have selected taxonomically and perceptually similar items, comforting the idea that shape similarities are important in their decision process (Kucker et al., 2019). This is interesting and important, as it shows that comparisons increase sensitivity and reduces the bias with respect to no-comparison conditions.

These results lead us to conclude first that simultaneous availability helped children generalize and that in this condition children are both discriminating well between generalization items and are unbiased. This confirms our prediction for response bias and reveals that simultaneous availability does improve sensitivity a question we had left open. In the light of these results,

simultaneous availability appears a powerful factor that can promote conceptually based generalization by improving both children's sensitivity and strategies thanks to the multiple comparisons between learning and generalization this condition allows.

Second, we can consider the sequential condition in which children's sensitivity is lower, children's strategies are more conservative, and both indexes are affected by learning distance. The main difference between learning conditions is due to the proportion of false alarms (see Table 2). In sequential availability and close learning children make few false alarms. It is in this crossing of conditions that children are the most conservative.

It is only in this availability condition that we can consider the effect of learning distance. We predicted that a far learning distance would improve children's sensitivity because such a learning situation should favor a broader category representation and word extension to higher category level members like the taxonomically related generalization items given here. However, in the present results, far learning – and the broader representation that it should enable – helps children be less conservative (i.e., far learning reduces the response bias) but it is close learning situations that improve sensitivity not far learning situations. Only a free-choice design could reveal that conceptual distance between items, can have this diverging effect on children's sensitivity and response strategy.

These result help us rethink the debate about the effect of amount of category knowledge and item's category level in the domain of novel noun generalization, and children's ability to extend novel words to items beyond the category's basic level (Jenkins et al., 2015; Xu & Tenenbaum, 2007). The stumbling block in this debate is that extra category knowledge from items at the category's subordinate level has opposite effects in different studies: it either promotes narrow generalization (Xu & Tenenbaum, 2007) or broader generalization (Jenkins et al., 2015). In the present study, both amount of knowledge and item's category level are manipulated, even if this is in a slightly different way, and an increase in the amount of knowledge available about the category (in simultaneous availability vs sequential or comparison vs no comparison situations) improves both sensitivity and generalization strategy. But conceptual distance (far learning) can improve strategies while reducing sensitivity which may be the source of what seems to be conflicting evidence in the previously cited papers. This debate is in the case of forced-choice designs and we think a free-choice approach, in which one can analyze sensitivity and strategies independently, may help find an outcome.

Finally, it is rather surprising to notice that the sequential condition – which of the two availability conditions may be the closest to a daily life situation

– is the condition in which children are highly conservative. This could mean that children in daily life situations may not name items for which they notice a conceptual relation because their strategy is conservative and would need further investigation to be confirmed.

Acknowledgments

We wish to thank the french Agence Nationale pour la Recherche (ANR) for their financial support in the COMPARE project to the third author. The first author benefited from a French ministry scholarship for research. We also wish to thank Zoé Mandernach for her help in the stimulus construction and in collecting data.

References

- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons : A Meta-Analytic Review. *Educational Psychologist*, 48(2), 87-113. <https://doi.org/10.1080/00461520.2013.775712>
- Gentner, D., & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition*, 2(2), 261-283. <https://doi.org/10.1515/langcog.2010.011>
- Gentner, D., & Namy, L. L. (1999). Comparison in the Development of Categories. *Cognitive Development*, 14(4), 487-513. [https://doi.org/10.1016/S0885-2014\(99\)00016-7](https://doi.org/10.1016/S0885-2014(99)00016-7)
- Jenkins, G. W., Samuelson, L. K., Smith, J. R., & Spencer, J. P. (2015). Non-Bayesian Noun Generalization in 3- to 5-Year-Old Children : Probing the Role of Prior Knowledge in the Suspicious Coincidence Effect. *Cognitive Science*, 39(2), 268-306. <https://doi.org/10.1111/cogs.12135>
- Kucker, S. C., Samuelson, L., Perry, L., Yoshida, H., Colunga, E., Lorenz, M., & Smith, L. (2019). Reproducibility and a Unifying Explanation : Lessons from the Shape Bias. *Infant behavior & development*, 54, 156-165. <https://doi.org/10.1016/j.infbeh.2018.09.011>
- Landau, B., Smith, B., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 299-321.
- Markman, E. M. (1989). *Categorization and Naming in Children : Problems of Induction*. MIT Press.
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
- Smiley, S. S., & Brown, A. L. (1979). Conceptual preference for thematic or taxonomic relations : A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology*, 28(2), 249-257.
- Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, 114(2), 245-272. <https://doi.org/10.1037/0033-295X.114.2.245>

**Chapter 3 - Novel noun generalization in a free-choice design:
investigating generalization constraints.**

The following chapter contains the proceedings paper: Stansbury E., Witt A., Thibaut J-P., (2023). Novel noun generalization in a free-choice design: investigating constraints. *Proceedings of the Annual Meeting of the Cognitive Science Society, 45th*. In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), pp. (3345-3351).

Note: Amendments made to the paper for the clarity of the thesis are in boxes.

Novel Noun Generalization in a Free-Choice Design : Investigating Generalization Constraints.

Eleanor Stansbury (eleanor.stansbury@u-bourgogne.fr)

Arnaud Witt (arnaud.witt@u-bourgogne.fr)

Jean-Pierre Thibaut (jean-pierre.thibaut@u-bourgogne.fr)

LEAD-CNRS, UMR 5022, Université de Bourgogne Franche-Comté,
Pôle AAFE – Esplanade Erasme, 21065 Dijon, France

Abstract

A common result in novel word generalization is that comparison settings (i.e., several stimuli introduced simultaneously) favor taxonomically-based generalization. Most generalization studies on comparison have been done with forced-choice designs. We investigated which type of items five-year-old children would choose as referents in a free-choice novel noun generalization task. Options were items from the same basic level category, from a near superordinate category, a distant superordinate category, and also perceptual lures, thematic lures, and unrelated lures. We manipulated the generalization items availability at test (i.e., generalization stimuli introduced sequentially or simultaneously). Results show that items from the same basic level category were more chosen than other taxonomically related items. Interestingly, perceptual lures and near superordinate items did not differ, suggesting that children did not arbitrate between perception and taxonomy. Results are discussed in terms of the respective role of taxonomic relations and perception but also mode of presentation (availability).

Keywords: Categorization, generalization, novel noun, forced-choice, free-choice, conceptual distance, stimuli availability

Introduction

Comparisons and novel word generalization

Given the importance of language learning, identifying which word learning situation promote concept construction and novel word generalization is an important topic for cognitive sciences. The available evidence during learning constrains children's later generalization. For example, being introduced to one learning stimulus or to several stimuli does not provide the same amount of evidence and thus do not allow to test hypotheses regarding a novel noun extension with the same amount of certainty (e.g., Markman, 1989). One prevailing view is to consider that novel noun generalization is biased towards a limited number of dimensions (Markman, 1990). In experimental design involving one single stimulus, various biases have been described. They show that children tend to give the same name to objects that are holistically similar, or to objects that display the same shape or to objects of the same basic level category (Emberson et al., 2019; Imai et al., 1994; Jones &

Smith, 1993; Landau et al., 1988b; Markman, 1989a; Markman & Hutchinson, 1984; Waxman, 1990; Waxman et al., 1991).

One way to reduce these so-called biases might be to introduce several learning instances. Indeed, the opportunity to compare several learning exemplars for a novel word would favor conceptually based categorization, compared to the classical one-single exemplar learning design (Childers, 2020; Gentner & Christie, 2010). This is the case because, introducing several items decreases the number of possible dimensions to which the noun may be generalized. Thus, the comparison process constrains which dimensions are more plausible. In a novel word learning by comparison design, two or more learning items are displayed (e.g., two apples, or two fruits) and belong to the same taxonomic category. In the generalization phase, the child has to choose between an item that is taxonomically-related to the learning items but rated as perceptually dissimilar to them (e.g., a banana) and a perceptually similar lure that is conceptually unrelated to the learning items but perceptually similar (e.g., a red Christmas ball) to them (Gentner & Namy, 1999). Whatever the variations and particulars of the design, studies have shown that comparison situations and the presence of a unifying name (i.e. label effect) lead to more conceptually-based generalizations than no-comparison (single stimulus) situations.

However, when one compares the available studies using the comparison design, depending on the comparison condition implemented in the experiment, the percentage of taxonomic choices varies across studies (from 50% to 80% and beyond). These differences suggest that various factors might modulate the effectiveness of comparison learning situations.

One factor might be the format of the experiment. Most existing studies on novel object noun learning and generalization use forced-choice designs (Alfieri et al., 2013) This format has been used to study which commonalities children spontaneously prefer in novel noun learning (e.g., biases in novel word generalization). A limited set of options that are

hypothetically potential targets for generalization (e.g., same shape, versus same texture, versus same color) (Landau et al., 1988a; Markman, 1989b). Children have to choose “one that shares the same name” among the options.

As a result, they might select an item that they would not have selected as an item of the same category if they considered that none belong to the same category. Conversely, selecting one option does not mean that participants would not accept the other options as members of the category, as shown by Smiley and Brown (1979). In order to allow children to choose as many (or as little) items when generalizing nouns we use a free choice design. We also used a comparison design (see hereafter).

Comparison as alignment

It has been claimed that showing two or more stimuli simultaneously invites comparison of the stimuli and would contribute to elicit a different, more relevant, conceptual encoding. Comparing the learning items promotes the alignment of their common features (Gentner et al., 1999) and labels (Namy & Gentner, 2002). The first output of a comparison process would be composed of common salient perceptual features. These common features would promote more comparisons that would elicit further alignments based on less salient conceptual commonalities (Namy & Gentner, 2002). The representation of both objects resulting from this stimulus mapping process is more conceptually based than the one that would have been built without the possibility to compare stimuli, that is in a single object design.

Here we focus on several factors which might potentially influence comparison. Semantic distance is one such potential factor, as it has been shown that children have more difficulties conceptualizing superordinate categories than basic level categories (Rosch et al., 1976). On the other hand, data that assess the role of distance between items in comparison designs is scarce (Liu et al., 2001). Thibaut and Witt (2017) manipulated the semantic distance between learning items and showed that comparing learning pairs from more distant domains led to better taxonomic generalization for children from both age groups.

No comparison designs tell us that there is a basic level bias, according to which young children spontaneously extend a novel object to items from the same basic level of categorization rather than to other levels of categorization (e.g., *apple* from one apple to any other apple rather than *Cox apple* or *fruit*) (Markman, 1989; Waxman et al., 1991; Waxman & Hatch, 1992). This bias can also explain children’s difficulties with superordinate categories in the comparison case.

Goals of the present experiment

In two experiments, we studied 5-year-old children in a comparison design. The difference with previous studies is that we use a free-choice design rather than a forced-choice design, in which we will introduce two learning exemplars and different categories of options.

Most existing studies on object noun comparison and generalization use forced-choice designs (Alfieri et al., 2013b). They have been used to study which commonalities children spontaneously prefer in novel noun learning (e.g., biases in novel word generalization). They introduce a limited set of options that are hypothetically potential targets for generalization (e.g., same shape, versus same texture, versus same color) (Landau et al., 1988a; Markman, 1989b). Most often, children have to choose only one item among the options (e.g., the one that shares the same name).

As a result, they might select an item that they would not have selected as an item of the same category in a free choice design, for example if they considered that none belong to the same category. Conversely, selecting one option does not mean that participants would not accept the other options as members of the category. For example, Smiley and Brown (1979) showed that young children could select and justify a taxonomic choice after a first, thematic, choice, when given the opportunity to make a second choice. It is to overcome this limit of forced choice designs and allow children to choose as many (or as little) items when generalizing nouns that we use a free choice design.

We also manipulate semantic distance at learning and between learning items and options, and the type of lures. The taxonomically related options will belong to the same basic level category or the same immediate superordinate category (e.g. banana, for two learning apples) or the same distant taxonomic category (e.g., meat for the two apples). The lures will be perceptually related or thematically related or unrelated with the learning items. Contrasting different taxonomic distances is important, and assesses whether children will choose all the taxonomically related items or choose them less as the semantic distance decreases. Another question is whether choosing taxonomically related items will eliminate perceptual and/or thematic choices. In forced-choice tasks, they are handled as incompatible options.

Testing the role of semantic distance between learning items is important in order to assess breadth of the category representation. Based on existing studies, we hypothesized that basic level categories would be selected more often than other taxonomically related items (see Thibaut & Witt 2017; submitted).

We also manipulated what we called item’s availability at test. Are all options introduced simultaneously at test, or are they introduced one by one? In daily life, children might see new

(generalization) items simultaneously (e.g., a fruit with other target fruits, or a fruit in a kitchen with other fruit-related objects, all on a table, or any display). They might also see the same generalization items one by one, sequentially, in a book, or introduced one by one. Those two modes of presentation might have different consequences. For example, children might compare the taxonomic options in the simultaneous case and include more distant taxonomic choices, by progressive alignment, than in the sequential case.

One important and open question is whether children will choose perceptual lures. Their selection might be modulated by the mode of presentation of the options: for example, perceptual lures could be chosen more often when stimuli are introduced one by one (sequential) because they cannot compare them with the taxonomic choices.

The second experiment is a replication of the first experiment except that we removed the two options from the same basic level category. We conjectured that children in Experiment 1 might select less same superordinate items as a default strategy or as a security strategy. Given that the instructions are to find items that share the same name, children might select the basic level options first and might be more reluctant to select less similar items or simply stop because they did the job.

Experiment 1

Methods

Participants

Seventy-nine French speaking children (forty-five boys) were tested individually in a quiet room at their school. Children were from middle class town center schools and were five-years-old (mean age: 54 months; 45 – 66). Children were randomly assigned to one of the two availability conditions (sequential: 40 children; simultaneous: 39 children).

The procedure was in accordance with the declaration of Helsinki (1964) and was ethically reviewed and approved within an Official agreement (convention n°: 2019-0679 and endorsement n°: 2020-0566) between the Academia Inspection of the French National Education Ministry, the University of Bourgogne (“Inspection Académique de Côte d’Or”). Participant’s consent was assured by an information letter sent to children’s parents and their returned written consent.

Materials

Color pictures of real objects were used as stimuli. The pictures were organized into sixteen stimulus sets, each associated with a semantic category (e.g., accessories, foods, clothing, tools, etc), each set was designed with three learning stimuli and twelve generalization stimuli. The sixteen trials were divided into two learning conditions (close or far

learning). Each trial was constructed around a semantic category. In each learning condition, one of the two pictures was considered as the standard picture. In the close learning condition, the two learning items were two pictures of objects from the same basic level category (e.g. a pear and a cut pear). In the far learning condition, the two learning items were from the same superordinate category (e.g., a pear and a raspberry).

The twelve generalization items were: two pictures of objects from the same basic level category as the standard learning item (basic level category generalization, TaB, e.g., pears), two pictures from the same superordinate level category as the learning-items (near superordinate generalization items, TaN, e.g., apricots and pineapple), two pictures from a more distant superordinate category as the learning-items (distant superordinate generalization items, TaD, e.g., chips and pasta); two stimuli perceptually similar to the standard learning item but not taxonomically related to the learning items (perceptual distractors, P, e.g. a punching ball and a pear shaped candle); two pictures thematically but not taxonomically related to the learning-items (thematically related distractors, Th, e.g., a fruit basket and a fruit knife); two lures semantically and perceptually unrelated to the learning items (non-related distractors, NR, e.g. a car and a note book).

The trials’ order during the task was balanced. All 16 trials in a task were presented in one availability level (simultaneous or sequential). In the sequential generalization availability, the generalization-items’ order was balanced between trials. In the simultaneous generalization availability, the position of the generalization items on the screen was balanced between trials. Figure 1 shows an example of a trial built using the stimuli from the food/fruit/pear category. The pictures were displayed on a 13inch touchscreen laptop.

We forged 16 different bisyllabic labels (pseudo-words) which are, as shown by Gathercole and Baddeley (1993), easier to remember than monosyllabic pseudo-words (e.g., buxi, dajo, zatu, xanto, vira). Syllables were of the CV type which is the dominant word structure in French (from Lexique.org, New, Pallier, Brysbaert, & Ferrand, 2004).

Ratings for generalization items. Twenty-eight students’ ratings confirmed that taxonomically related items were considered to belong to the same category as the standard learning item (average ratings: $M_{Ta} = 7.6$, $M_{Th} = 4.5$, $M_P = 2.1$, $M_{NR} = 1.5$, average p between Ta-Th $p < .001$). Twenty-four students rated near taxonomically related generalization items conceptually closer to the standard item compared to distant taxonomically related items.

Perceptual similarity ratings were obtained from 36 students and controlled that perceptual matches

were perceptually more similar to the learning items than all other generalization items (average ratings: $M_{Ta} = 3.0$, $M_{Th} = 2.2$, $M_P = 6.3$, $M_{NR} = 1.7$, average p between P-Ta $p < .001$). Ratings were also obtained from 21 students for thematically related distractors. They showed that they were more thematically related to the learning items than all other generalization items (average ratings: $M_{Ta} = 6.4$, $M_{Th} = 7.5$, $M_P = 2.2$, $M_{NR} = 1.8$, average p between Th-Ta $p < .05$). Unrelated distractors scored significantly below all other generalization items in all ratings ($p < .01$)

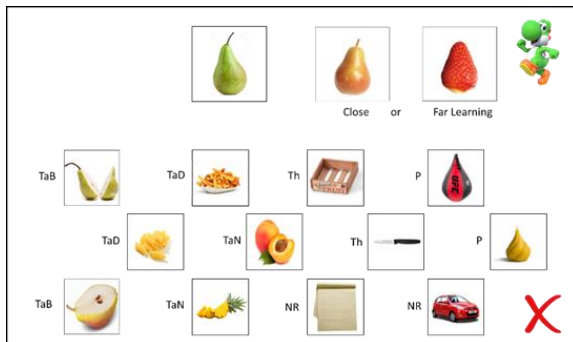


Figure 1: Trial built for the food category

Note: Participant saw either the close or the far learning item

TaB : same basic taxonomic category, TaN : near superordinate category, TaD: distant superordinate, Th: thematically related, P: perceptually related, NR : non related generalization items.

Procedure

Participants were seated at a low table, in a quiet room at their school, facing the laptop, next to the experimenter. They were randomly assigned to one of the generalization availability conditions (sequential, or simultaneous). In both conditions, children were introduced to a puppet named “This is Yoshi, we are going to play with him. But he lives far away from here and speaks a different language. In the game we are going to learn his language.” The experimenter then showed the fifteen trials. In all two learning conditions learning items appeared one by one near the top of the screen and the experimenter announced their name as they appeared using the instruction: “Yoshi’s mummy says that this is a *buxi*, and this one is also a *buxi*; Yoshi must find other *buxis* for his mummy...”. Then, the generalization items appeared on the lower part of the screen, generalization availability one by one in the sequential condition, the experimenter said “is this a “*buxi*...?” for each of the 10 generalization items. In the simultaneous condition, they were displayed simultaneously: “which ones of these are also *buxis*, show me the *buxis* but not the other things”. The experimenter finished the instructions by “Take your time, don’t give me your answer before Yoshi appears on the screen”.

Design

Availability (sequential, simultaneous) a between-subject factor was crossed with learning distance (close, far) and generalization item (basic, near superordinate, distant superordinate, theme, perceptual, non-related) within-subject factors.

Results

We recorded the number items chosen by children in each trial for each category of generalization item (i.e., basic category level, near superordinate, distant superordinate, perceptually related, thematically related, non-related) and the order in which items were chosen. The proportion of answers for each type of generalization item was calculated out of the total number of items of each type across all 16 trials (i.e., 32 items). This means that all or no items could be chosen in a trial, and that any type of generalization item could be chosen at any order in the sequence of choices.

Children’s proportion of answers analysis

We ran a three-way repeated measures ANOVA on the proportion of answers with availability (sequential, simultaneous) as a between factor and learning distance (close, far) and generalization items (basic level, near superordinate, distant superordinate, theme related, perceptual match, non-related) as between factors.

Results reveal significant differences in children’s proportion of choices of generalization items $F(5,385) = 155.39$, $p < .001$, $\eta^2_p = .67$ (Figure 2). Student t test contrasts show that children choose more basic level category items than all other items ($p < .001$, Bonferroni corrected significance level at $p = 0.005$). They also selected more near superordinate items than thematically related and unrelated items ($M_{Near} = 0.40$; $M_{Theme} = 0.18$; $M_{Unrelated} = 0.16$; $p < .001$). The number of selections of near superordinate items did not differ from the selections of perceptual matches ($M_{Perceptual} = 0.42$; $p = 0.08$).

This analysis also revealed two interactions, between learning distance and generalization items $F(5,385) = 2.89$, $p < .05$, $\eta^2_p = .04$, and between availability, learning distance, and generalization items $F(5,385) = 3.15$, $p < .01$, $\eta^2_p = .04$. In this latter interaction, we were mostly interested in difference between the two availability levels across the other levels of the two other variables. Student t test contrasts with a Bonferroni corrected significance level set at $p = .0011$, surprisingly, revealed no difference between the two availability levels across all the levels of the two other variables ($p > 0.05$).

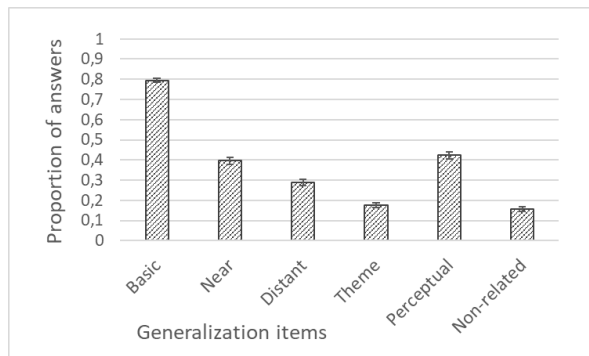


Figure 2: Proportion of answers in a 12 item task as a function of generalization item (basic level, near superordinate, distant superordinate, thematically-related match, perceptual match, non-related). Error bars a SEM.

Children's answers order analysis

We ran a two-way repeated measures ANOVA on average order in which items had been chosen with learning distance (close, far) and generalization items (basic level, near superordinate, distant superordinate, theme related, perceptual match, non-related) as between factors.

Results revealed an effect of generalization item $F(5,95) = 12.40, p < .001, \eta^2_p = .22$ ($M_{Basic} = 2.83$; $M_{Near} = 4.19$; $M_{Distant} = 4.50$; $M_{Theme} = 3.66$; $M_{Perceptual} = 4.02$; $M_{Non-related} = 5.31$). Children chose basic category level items before the other generalization items except for thematically related items ($p < .01$) which confirms children's tendency to choose the basic category level items with certainty and to find the other taxonomically related items and perceptual lures more ambiguous.

Addition to the paper:

Signal detection theory analysis

For the data analysis we calculated a sensitivity index D' and a response bias index β derived from signal detection theory (Macmillan & Creelman, 2004) adapting them to experiments based on small numbers of stimuli (see, Rioux et al., 2018b). D' (range 0 : 1) indicates participants ability to discriminate (high values for better discrimination) and β indicates their strategy (liberal, for negative values of β ; conservative for positive values of β ; neutral for $\beta=0$).

Sensitivity We ran a two-way repeated measure ANOVA on the sensitivity index D' with availability (sequential and simultaneous) as a between-subject factor and learning distance (close and far) as a within factor. Results did not reveal any significant effects of these factors ($M_{Sequential} = 0.70, M_{Simultaneous} = 0.60$).

Response bias We ran a two-way repeated measure ANOVA on the response bias index β with availability (sequential and simultaneous) as a between-subject factor and learning distance (close and far) as a within factor. Results did not reveal any

significant effects of these factors ($M_{Sequential} = 0.27, M_{Simultaneous} = 0.23$).

Discussion

This experiments' main aim was to compare the number of different types of stimuli in a comparison format, as a function of stimulus availability and semantic distance between learning items. The main result was that children chose more items from the same basic level category than any other types of available generalization options. However interestingly, in both availability conditions, they chose a significant number of stimuli from the near superordinate category and perceptual lures at the same level. This means that beyond basic level objects, they also accept items that are taxonomically related at a higher category level and items that are perceptually related. They do not arbitrate between taxonomic relations and perception. We will come back to this in the general discussion. Also of interest, availability played no role while generalization performances' were affected by this factor in previous studies (Stansbury et al., 2022).

Addition to the paper:

Signal detection indexes were not affected by availability either, unlike in previous studies. The presence of same basic category level items, that are significantly more chosen than all other items, seem to equate both availability conditions.

Experiment 2

Methods

Participants

Fifty French speaking children (twenty-nine boys) were tested. Children (mean age: 16 months; 50 – 62) were from middle class town center schools. Children were randomly assigned to one of the two availability conditions (sequential: 18 children; simultaneous: 32 children). The ethical procedures were identical to experiment 1.

Materials

The same material as in experiment was used except that the same basic level options were removed, for a total of 10 options. All other aspects of materials were identical to experiment 1.

Procedure

The procedure was identical to the one followed in Experiment 1.

Results

Proportion of choice analysis

We ran a three-way repeated measures ANOVA on the proportion of answers with availability (sequential, simultaneous) as a between-subject factor and learning distance (close, far) and generalization item (near superordinate, distant superordinate, theme related, perceptual match, non-related) as within-subject factors.

Results revealed that children choose a higher proportion of items in simultaneous settings rather than sequential settings $F(1,55) = 6.63, p < .05, \eta^2_p = .11$ ($M_{\text{close}} = 0.27; M_{\text{far}} = 0.29$). Children also choose more items in far rather than close learning settings $F(1,55) = 11.2, p < .001, \eta^2_p = .17$ ($M_{\text{sequential}} = 0.18; M_{\text{simultaneous}} = 0.37$). Different generalization items are also chosen in significantly different proportions $F(4,220) = 24.46, p < .001, \eta^2_p = .32$: children choose more near superordinate items than distant superordinate items, thematically related lures and non-related items ($p < .001$, Student t tests, with Bonferroni corrected significance level at $p = 0.005$), however they choose as many near superordinate items as perceptual lures.

This analysis also revealed an interaction between learning distance and generalization item $F(1,55) = 6.63, p < .05, \eta^2_p = .11$ (Figure 3). In this interaction the differences of interest are those between close and far learning cases for same generalization item types. Interestingly, the only significant difference between the two learning conditions was observed for perceptual matches ($p < .001$, Bonferroni corrected significance level at $p = 0.005$), that are chosen more in far rather than close learning settings ($M_{\text{close}} = 0.31; M_{\text{far}} = 0.41$).

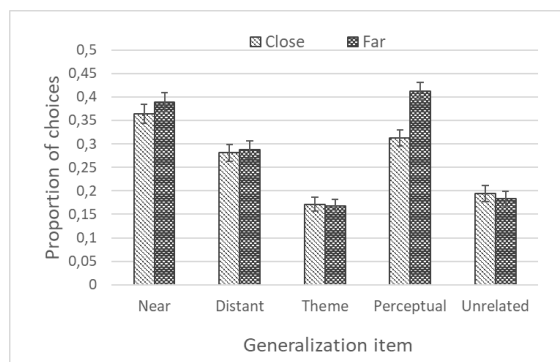


Figure 3: Proportion of answers in a 10-item task as a function of learning distance (close, far) and generalization item (near superordinate, distant superordinate, thematically-related, perceptual match, unrelated). Error bars a SEM.

Addition to the paper:

Signal detection theory analyses

For the data analysis we calculated a sensitivity index D' and a response bias index β derived from signal detection theory (Macmillan & Creelman,

2004) adapting them to experiments based on small numbers of stimuli (see, Rioux et al., 2018b). D' (range 0 : 1) indicates participants ability to discriminate (high values for better discrimination) and β indicates their strategy (liberal, for negative values of β ; conservative for positive values of β ; neutral for $\beta=0$).

Sensitivity We ran a two-way repeated measure ANOVA on the sensitivity index D' with availability (sequential and simultaneous) as a between-subject factor and learning distance (close and far) as a within factor.

Results didn't reveal any significant effect. It is interesting to note that sensitivity levels are relatively low ($M_{\text{Sequential}} = 0.25, M_{\text{Simultaneous}} = 0.25$).

Response bias We ran a two-way repeated measure ANOVA on the response bias index Beta with availability (sequential and simultaneous) as a between-subject factor and learning distance (close and far) as a within factor.

Results revealed a simple effects of availability $F(1,55) = 13.20, p < .001, \eta^2_p = .19$ where children's response bias is more conservative in sequential setting compared to simultaneous settings and learning $F(1,55) = 7.27, p < .01, \eta^2_p = .12$ where the response bias is more conservative after close learning than after far learning ($M_{\text{Close}} = 0.70; M_{\text{Far}} = 0.59$).

Comparison between the two experiments.

Proportion of choice analysis

We ran a four-way repeated measures ANOVA on the proportion of choices with generalization items (near and distant superordinate, thematically related, perceptual match, unrelated) as a within-subject factor, availability, learning distance, Experiment (1 and 2) as between-subject factors. Here only the differences between experiments and between generalization items are of interest. Results revealed an interaction between the experiment (i.e., 12 or 10 item experiment) and availability $F(1,132) = 5.46, p < .05, \eta^2_p = .04$ (Figure 4). Student t test contrasts (Bonferroni corrected significance level $p = 0.006$) reveal that children make more choices in simultaneous rather than sequential settings ($p < .001$) and this is the case in experiment 2 only.

The analysis also revealed a four way interaction between number of items, availability, learning distance and generalization $F(4,528) = 3.90, p < .01, \eta^2_p = .03$. The differences between group were small. Indeed, a posteriori Student t test contrasts, with Bonferroni correction, revealed no significant difference between Experiments 1 and 2 for any generalization item in any experimental condition.

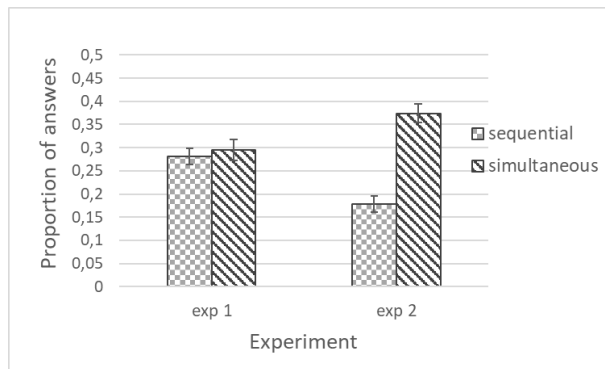


Figure 4: Proportion of answers a function of experiment (1, 2) and availability (sequential, simultaneous). Error bars are SEM.

Addition to the paper:

Signal detection theory analyses

Sensitivity We ran a two-way ANOVA on the sensitivity index D' with experiment (1, 2) and availability (sequential and simultaneous) as between-subject factors.

Results revealed an effect of experiment $F(1,132) = 5.84, p < .01, \eta_p^2 = .07$. Children sensitivity is significantly lower in experiment 2 where 10 generalization items were shown than in experiment 1 with the two additional basic category level items ($M_1 = 0.65, M_2 = 0.23$).

Response bias We ran a two-way ANOVA on the response bias index Beta with experiment (1, 2) and availability (sequential and simultaneous) as between-subject factors.

Results revealed effects of all factors, availability $F(1,132) = 12.26, p < .001, \eta_p^2 = .09$, experiment $F(1,132) = 4.30, p < .05, \eta_p^2 = .03$ and an interaction of both factors $F(1,132) = 6.54, p < .01, \eta_p^2 = .05$. The interaction between experiment and availability revealed that while the response bias is similar in both experiments in simultaneous settings (simultaneous : $M_1 = 0.37, M_2 = 0.31$) and is close to a neutral bias, in sequential settings the response bias is significantly more conservative in experiment 2 with 10 generalization items (sequential : $M_1 = 0.47, M_2 = 0.99$).

Discussion

We studied children's word extension in a comparison setting with a free-choice design rather than the more common forced-choice word generalization design. Indeed, a free-choice design allows children to choose different categories of stimuli, including the lures, as items bearing the same noun as the learning items. Hence, they were able to choose both taxonomically related and perceptual (or thematic) lures. This is not possible with forced-choice designs pitting taxonomic choices and perceptual (or thematic) lures. Of central

interest was the number of selections of each type of options as a function of conceptual distance and mode of presentation of these options (items' availability).

Both experiments tested which taxonomic options children would choose. They did not select the three taxonomic distances in the same way. The same was true for the lures. In Experiment 1, children selected the basic level options much more frequently than the two other taxonomically related items. This is consistent with previous research on lexical biases showing, in the case of no comparison designs that children are biased towards the basic level of categorization (Markman & Hutchinson, 1984; Rosch et al., 1976). Importantly, this result shows that despite this basic level bias, they could also accept taxonomically more distant items. In contrast, these results did not confirm an overall preference for *any* other taxonomically related choice (near or distant superordinate) (Waxman, 1990).

What our data adds to the existing literature is that our free-choice comparison design led to generalizations influenced by multiple biases at the same time. What was less predicted is that distance between the learning items had no impact on these taxonomic choices. This does not seem consistent with existing results obtained with a recent Bayesian approach, hypothesizing that more distant learning items would increase the number of choices of superordinate category items (Xu & Tenenbaum, 2007a, 2007b). With learning objects from two different basic level categories, one would have expected more generalizations of the nouns beyond the basic level categories, or smaller differences between the number of basic level and superordinate level choices in the far learning condition. This was not the case.

Another important aspect of the experiment is the choice of lures. As expected, unrelated lures were marginally chosen. So were the thematically related lures. This latter result confirms previous results showing that the interpretation of nouns is biased towards taxonomic relations rather than other semantic relations (Waxman & Kosowski, 1990). As for the perceptual lures, their status is more ambiguous. Single designs are correlated with same shape choices, like our perceptual lures (Jones et al., 1991; Kucker et al., 2019).

Interestingly perceptual lures were often chosen as much as near superordinate level choices and more often than distant superordinate choices. This interesting result would have been hidden in forced-choice designs as, in these designs, only one option (perceptual or taxonomic) can win. Our results suggest that taxonomic choices do not eliminate perceptual choices and are consistent with the enduring influence of perceptual similarities.

A final result is that the availability of the items (sequential or simultaneous) had no important impact in Experiment 1 but resulted in significantly

less choices in Experiment 2. This is probably due to the fact that children chose a high percentage of the basic level items in both conditions in Experiment 1, which contributed to equate the two conditions. The presence of these basic level objects might have elicited a less cautious strategy in the sequential case, because, with these obvious choices, children were more eager to select stimuli. This remains to be tested. The comparison of the two experiments, on equivalent stimuli, is consistent with this analysis as there was no difference between the two conditions in Experiment 1 whereas the difference is important in Experiment 2 (as shown by the interaction between availability and Experiment).

Concluding remarks.

Our experiments showed that children could select different types of taxonomically related items in the same set of options, but the proportions of these choices could vary as a function of distance. These taxonomic choices were paralleled by perceptual

lures choices, a result which would have remained hidden in a classical forced choice. This suggests that the extension of the nouns incorporate different types of information in young children. It will be interesting to see whether older children or adults would no longer hesitate between these two sources of information.

Acknowledgment

We wish to thank the French Agence Nationale pour la Recherche (ANR) for their financial support for the COMPARE project (project ANR-18-CE28-0019), to the third author. The first author benefited from a scholarship of the French ministry of Higher Education and research.

References

La mise à jour automatique des citations est désactivée. Pour voir la bibliographie, cliquez sur Actualiser dans l'onglet Zotero.

GENERAL DISCUSSION

In the empirical work presented in 6 of the 7 chapters of the present thesis, our main question was: *how* do children come to a conceptual generalization decision? Our aim was to investigate the strategies children use to generalize novel nouns on a conceptual basis and investigate the processes that underlie such generalizations.

One of the major challenges for children learning words is understanding a word's meaning (Bloom, 2002). Cognitively this implies building the right concept – i.e. mental representation with weighted features (see General Introduction) - for the right word. In our work, with novel noun generalization tasks, children's challenge was to build a conceptually based mental representation that would enable them to make a conceptually based generalization of a novel noun. Consequently, a first point of interest is what our work says of how children build their representations when working towards a conceptual generalization.

Building a concept in novel noun generalization tasks: the role of learning and generalization items.

Our two eye tracking analyses allowed us to describe clear patterns of novel object noun generalization and novel relational noun generalization. To the best of our understanding this is the first time that novel noun generalization strategy profiles and their temporal dynamics have been described. Both studies revealed a consistent strategy profile across age groups: an Elimination or Alignment type strategy (Bethell-Fox et al., 1984; Markman & Gentner, 1993) for object noun generalization and a Construction or Projection (Bethell-Fox et al., 1984; Hummel & Holyoak, 1997) type strategy for relational noun generalization.

Numerous studies have outlined comparisons between learning exemplars as an important process underlying word learning – and thus representation building – before word generalization (Alfieri et al., 2013). The suggestion behind this is that, following the structural alignment hypothesis (Gentner & Colhoun, 2010; Gentner & Namy, 1999, 2004; Markman & Gentner, 1993; Thibaut & Witt, 2023), comparisons between same category members highlight their conceptual features that may have otherwise gone unnoticed. Thus, comparison is a fundamental process for building representations with conceptual features. Both Elimination and Construction type strategies were consistent with early comparisons between learning items and confirmed that learning items, and comparisons between them, play an anchoring role in novel noun generalization.

A new line of reflexion that was raised by the strategies revealed in the eye tracking analyses, and was then addressed in Section 3, was the impact of generalization items on the representation built during novel noun generalization tasks. In both object and relation noun generalization, learning items and generalization items were considered simultaneously at some point during the task, suggesting that information from both types of items were being processed simultaneously. Given such a result, a compelling question was whether we were witnessing processes such as the re-representation processes suggested by Gentner & Colhoun, 2010, even though this was beyond the scope of the present studies. Eye tracking analyses also revealed strategy profiles that suggested, if not directed comparisons between generalization items, at least comparisons involving them (cf. learning item to generalization option matches in object noun generalization, and TargetEntity to generalization option in relational noun generalization). However, considering the powerful learning mechanics recognized in comparison, this result also suggested that generalization items may play a more important role in building the correct representation for generalization than it has usually been suspected.

Studies presented in Section 3 analysed children's generalization in free-choice designs, and results argued for the role of generalization items in children mental representation building. In these studies, generalization availability affected children's proportion of taxonomic choices, and more importantly, generalization availability affected both children's sensitivity (i.e. ability to discriminate same category members among the generalization options) and children's response strategy (i.e., willingness to use a taxonomic categorization rule). Children decide on category membership – discriminate between category members and non-members - if the items' features and the mental representation's features meet a good enough match. That generalization availability affected children's ability to discriminate category members is a strong argument for the role of generalization items on mental representations built and used for generalization.

Taken together our results confirm the anchoring role of comparisons between learning items for conceptual generalization. They also add to our understanding of representation building during generalization tasks by highlighting the role generalization items have in the process.

Developmental course in novel noun generalization

Our investigation of how children come to a conceptual generalization decision was in a developmental perspective. While we investigated children's process, we were interested in how they develop.

Surprisingly, the strategy profiles revealed in the eye tracking analyses were stable across age groups. We suggested that strategies used may develop with age and the development of children's executive functioning (Bethell-Fox et al., 1984; Lagarrigue & Thibaut, 2020). Strategy profiles were not subject to developmental change. Our investigation into the differences between age groups and the predictive factors of generalization success did however reveal a key to successful generalization: the strategies implementation. Children that focused longer on the learning items, fulfilling the comparison process, before directing their attention to the generalization items, were more successful. This may be due to children's better executive functioning, even though our investigation into executive functions did not confirm such and hypothesis. Interestingly, a clear relation was revealed between age and children's strategy implementation quality: the older the children were, the better they were at organizing their strategy execution. Let us highlight, that this is a consistent finding across generalization tasks even though the strategies revealed in both tasks were of different types.

Predictive power of early gazes in generalization strategies revealed that children need to focus on the learning items early on in the task and inhibit their inclination to consider the generalization items. This implies they resort to a top-down process to successfully build their mental representation of the targeted category (Sloutsky, 2010) and argues for the suggestion that from a certain age conceptual development is the result of both top-down and bottom-up process (see General Introduction, Ashby et al., 1998; Sloutsky & Fisher, 2004).

What may be questioned is what children are inhibiting? In the study of object noun generalization children's tendency to direct their attention to the generalization items before the learning items may be translated in terms either of the Goal Directed Strategy (Thibaut & French, 2016) or Smith and colleagues four step model (Smith et al., 2002). These two hypothesized strategy types have different implications. In the former children's top-down directed search for a goal must be inhibited, in the second the bottom-up attraction from the salience of the perceptive lure must be inhibited. We suggest that the attraction pulling children away from the learning items may evolve with age. Younger children's attention being strongly attracted to perceptive salience, slightly older children being guided by their

urge to find a solution, and only the children with a sufficient level of executive functioning being able to concentrate on learning items first. However, this is a suggestion, it is not demonstrated the present empirical evidence. What can be supported here, is the there is an interplay between top-down and bottom-up process in conceptual development. More importantly, we suggest that a developmental factor in conceptual developmental is the gradual change in balance between bottom-up and top-down processes: as more top-down abilities progressively come online, representation building and strategy implementation processes develop.

A categorical decision process at the end of generalization strategies?

Children's ability to focus on the learning items at the beginning of their search clearly appears as a predictive factor of successful conceptual generalization. Early attention to learning items guarantees sufficient information processing, sufficient comparisons between items, which would influence the features selected for the mental representation being built.

However, early attention orientation did not seem to be the only step related to generalization success. Decisional processes at the end of trials seemed to appear at the end of both object and relational noun generalization strategies. Indeed, it is only at the end of trials that children's attention focused on the item that they would then choose. Other argument of importance: it was only at the end of trials that differences between eye movements from trials that led to correct answers differed from those from trials that led to errors. While strategy profiles seemed to indicate that part of generalization success depended on the final decision children made between options, nothing in the eye tracking analyses enabled us to confirm this suggestion. More detailed results however came from the studies presented in Section 3 that analyzed children's generalization in free-choice designs.

The novel noun generalization tasks run with free choice designs enabled us to analyze two different components of children's generalization answers: on the one hand their sensitivity to category members or ability to discriminate category members between all available generalization options, and on the other hand, their response bias or willingness to answer, interpreted in this work as their willingness to follow a taxonomic categorization rule while generalizing. Children's sensitivity is tightly related to information processing and a relation between sensitivity and early attention at the onset of trials may be investigated in future research. Children's response bias however, may be the factor that can "swing" the decision process at the end of trials, somewhat independently of the search. Free choice

generalization studies revealed that children's response bias was affected by generalization item availability and conceptual distance between learning items. Overall, in these studies children's response bias was less conservative when all generalization items were available simultaneously, compared to when they were available sequentially. In other words, children were more willing to generalize the nouns to taxonomically related items when generalization items were presented simultaneously. This may be because more taxonomic information was available at once – due to the presence of all learning and all taxonomic generalization items at one - and cued children's decisions towards taxonomic items. Another interpretation could be that simultaneous generalization availability enables easy contrasts between generalization items with different statuses, that highlight that a decision must be made as to the generalization rule to follow. This may promote a decision process guided by stronger top-down processes than in a setting where generalization items are available sequentially. In this case, the children's higher performances would be explained by a better directed decision process. These results argue for a categorization decision step at the end of trials, swung by children's willingness to follow a taxonomic categorization rule. They also open the question as to what factors may improve children's willingness to categorize taxonomically: amount of taxonomical information at hand, or cues towards more complex decisional processes.

Further openings

Discussing strategy type differences in object and relational noun generalization

Two more points may be discussed before we conclude. First the strategy difference revealed between object noun and relational noun generalization strategies. Second the insights our results give on the effect of conceptual distance between items in novel noun generalization tasks.

Rather unexpectedly, eye tracking analyses revealed different strategy types for object and relational noun generalization. Thus, children's generalization strategies may be noun type dependent, or task dependent. This can be considered a limit to our work. However, it also leads to an interesting line of reasoning about children's development and the educational implications of our work. If generalization strategies are – at least to a certain degree – task dependent, then what is the relation between the task and the children's generalization performances? If an aspect of the task impacts children's generalization performances, which aspect is this and how does the task favor or impede novel noun generalizations?

To reflect on these questions, it is interesting to notice that children's relational noun generalization performances were rather high even though they were five years old, the youngest age we consider in the present eye tracking analyses. This is also noteworthy because relational nouns are known to be more difficult for children to learn than object names, and projection type strategies have been described as cognitively more demanding than elimination type strategies (Bethell-Fox et al., 1984).

It seems that children's performances seem less biased in the relational noun task than in the object noun generalization task. This may be due to less attractive lures in the relational noun generalization task than in the object noun generalization task that bias children's performances in a less strong way. In the novel object noun generalization tasks, the perceptive lure has very salient properties, mainly shape (Jones & Smith, 2002; Landau et al., 1988). However, in the relational noun generalization task one of the lures is a taxonomic match which likely is as attractive and entices as strong a lure. Thus, it may not be the lures status in relational noun generalization tasks that is responsible for children's high relational noun generalization performances.

We suggest that the particularly high scores in the relational noun generalization task may be explained by children's use of the projection type strategy and the way it constrains the generalization process to rely on the targeted relation's representation more than the task's concrete. Indeed, from a cognitive process point of view the main difference between the elimination and the projection type strategies, is that the first involves an elimination process during a concrete item to concrete item mapping process, while the second involves choosing the generalization solution while comparing the available options to a mental representation. Undoubtedly, the elimination type strategy tightly binds the cognitive processes to concrete items, whereas the projection type strategy involves more abstract processes, and constrains children to work on a more abstract level. Thus, could working with a projection type strategy make children less sensitive to biases? Future research would be needed to investigate this possibility. However, if it were the use of the projection type strategy in the relational noun generalization that favored conceptual generalization performances then research then it would become interesting in educational or remediation contexts to consider how tasks in themselves may constrain children's strategies and promote their generalizations.

Conceptual distance between novel noun generalization task items

Finally, we also manipulated the conceptual distance between items throughout the thesis's empirical work and found diverging results. Effects of conceptual distance rarely confirmed previously found results (Thibaut & Witt, 2015, 2017, 2023), and they were non consistent in the present studies. However, studies in free-choice designs analyzing both children's sensitivity and response bias in a signal detection framework may in light these contradictory results. Indeed, results showed that conceptual distance between learning items had reverse effects on sensitivity and response bias. A larger distance between learning items, reduced children's sensitivity to taxonomic items and impeded taxonomic generalization performances; while it made their response bias less conservative and favored taxonomic generalization performances. With these results at hand, that only a free-choice design and signal detection analyses could reveal, it is easier to understand how conceptual distant variations between items may have inconsistent effects.

Concluding remarks

The present work's developmental investigation into children's concept learning and generalization established a detailed description of children's novel noun generalization strategies. Children appeared to come to a conceptual generalization decision by means of specific generalization strategies, that were stable across development and dependent of noun type or generalization task. Empirical evidence highlighted both the importance of information processing for building a conceptually based mental representation and categorical decisional processes as underlying process to novel noun generalization. Our work confirmed the anchoring role of learning items and their comparisons for generalization, and added the suggestion that generalization items also play a role in the building of the mental representation that support the generalization processes. Finally, children's ability to control the implementation of their strategy appeared to be a key to successful conceptual generalization and the process subject to developmental change.

References

- Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning Through Case Comparisons : A Meta-Analytic Review. *Educational Psychologist*, 48(2), 87-113. <https://doi.org/10.1080/00461520.2013.775712>

- Ashby, F. G., Alfonso-Reese, L. A., Turken, A. U., & Waldron, E. M. (1998). A neuropsychological theory of multiple systems in category learning. *Psychological Review*, *105*, 442-481. <https://doi.org/10.1037/0033-295X.105.3.442>
- Bethell-Fox, C. E., Lohman, D. F., & Snow, R. E. (1984). Adaptive reasoning : Componential and eye movement analysis of geometric analogy performance. *Intelligence*, *8*(3), 205-238. [https://doi.org/10.1016/0160-2896\(84\)90009-6](https://doi.org/10.1016/0160-2896(84)90009-6)
- Bloom, P. (2002). *How children learn the meanings of words*. MIT press.
- Gentner, D., & Colhoun, J. (2010). Analogical Processes in Human Thinking and Learning. In B. Glatzeder, V. Goel, & A. Müller (Éds.), *Towards a Theory of Thinking* (p. 35-48). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03129-8_3
- Gentner, D., & Namy, L. L. (1999). Comparison in the Development of Categories. *Cognitive Development*, *14*(4), 487-513. [https://doi.org/10.1016/S0885-2014\(99\)00016-7](https://doi.org/10.1016/S0885-2014(99)00016-7)
- Gentner, D., & Namy, L. L. (2004). The Role of Comparison in Children's Early Word Learning. In *Weaving a lexicon* (p. 533-568). MIT Press.
- Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure : A theory of analogical access and mapping. *Psychological review*, *104*(3), 427.
- Jones, S. S., & Smith, L. B. (2002). *How Relevantproperties children know the relevant properties for generalizing object names*.
- Lagarrigue, Y., & Thibaut, J.-P. (2020). *From two to many : The role of executive functions in young children's generalization of novel object names in a comparison design*.
- Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development*, *3*(3), 299-321.
- Markman, A. B., & Gentner, D. (1993). Structural Alignment during Similarity Comparisons. *Cognitive Psychology*, *25*(4), 431-467. <https://doi.org/10.1006/cogp.1993.1011>
- Sloutsky, V. M. (2010). From perceptual categories to concepts : What develops? *Cognitive science*, *34*(7), 1244-1286.
- Sloutsky, V. M., & Fisher, A. V. (2004). Induction and Categorization in Young Children : A Similarity-Based Model. *Journal of Experimental Psychology: General*, *133*(2), 166-188. <https://doi.org/10.1037/0096-3445.133.2.166>

- Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological science*, *13*(1), 13-19.
- Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions : A developmental investigation with eye-tracking. *Cognitive Development*, *38*, 10-26.
- Thibaut, J.-P., & Witt, A. (2015). Young children's learning of relational categories : Multiple comparisons and their cognitive constraints. *Frontiers in Psychology*, *6*. <https://doi.org/10.3389/fpsyg.2015.00643>
- Thibaut, J.-P., & Witt, A. (2017). Generalizing novel names in comparison settings : Role of conceptual distance during learning and at test. *CogSci 2017 : 39th Annual Meeting of the Cognitive Science Society*, 3314-3319. <https://hal-univ-bourgogne.archives-ouvertes.fr/hal-01573610>
- Thibaut, J.-P., & Witt, A. (2023). Children's generalization of novel names in comparison settings : The role of semantic distance during learning and at test. *Journal of Experimental Child Psychology*, *234*, 105704. <https://doi.org/10.1016/j.jecp.2023.105704>