# MASTERARBEIT | MASTER'S THESIS

Titel | Title

The Effect of Referential Uncertainty on the Encoding of Novel Words


verfasst von | submitted by

Bojana Trajkovic


angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2024

| | |
|---|---|
| Studienkennzahl lt. Studienblatt \| Degree programme code as it appears on the student record sheet: | UA 066 013 |
| Studienrichtung lt. Studienblatt \| Degree programme as it appears on the student record sheet: | Masterstudium Joint-Degree Middle European interdisc. master prog. in Cognitive Science |
| Betreut von \| Supervisor: | Dr. Hanna Marno |

# Table of Contents

# Abstract

Learning a new word relies on processes that work at multiple time scales - learners need to identify a speaker's referent in-the-moment, encode a mapping between the label and the referent, recall multiple labeling events and integrate across them, and use their learned mappings to identify the object in novel contexts (Yurovsky et al., 2017). In the last decades, numerous research studies have addressed the question of "word-to-world" mappings and mechanisms responsible for the development of reference (Hollich et al., 2000). However, to date, no study has investigated whether and how subjective certainty about the word-referent relation influences learning and memory encoding process.

Therefore, this research study aims to extend the existing knowledge about word learning by investigating whether and how referential certainty might influence encoding processes during word learning. In particular, we test whether subjects' certainty about the meaning of a novel word would have an impact on the depth of encoding.

In order to investigate this question, we have recorded short movies (i.e., object naming episodes) where each time a person provided a novel word in the presence of two unfamiliar objects. Subjects' certainty about the meaning of the novel words was modulated by varying two factors: 1) object identity (by using two identical or two different unfamiliar objects) and 2) use of the referential cues (with having the speaker pointing at one of the objects or not during the labeling). After participating in the learning sessions, encoding of the novel words is assessed.

Our main hypothesis is that different levels of certainty about the meaning of the novel word will influence the depth of encoding. Specifically, we assume that the lower the uncertainty is about the meaning of the novel word, the more efficient word learning will be, and consequently, we expect encoding to be better.  Based on this assumption, we expect that encoding will be the best when the two different objects are presented with the person pointing towards one of them, and possibly also when two identical objects are shown in the absence of a pointing gesture, because in these conditions identifying the potential referent will be easy. In contrast, when the two identical objects are presented and there is a pointing gesture towards one of them, or when the two different objects are present in the absence of a pointing gesture, we

assume that uncertainty about word meaning will be high, and consequently, encoding will be worse than in the previous conditions.

We have conducted four experiments, three with both online and in-person versions. Based on our results, there are still insufficient grounds to claim that certainty about the meaning of incoming information would directly impact encoding novel words. Still, our results provide some interesting findings and an insightful look into the general relationship between memory encoding and certainty about the meaning of incoming information. Furthermore, the developed paradigm could be a good starting point for the later word learning research.

# Deutschsprachiges Abstract

Das Erlernen eines neuen Wortes beruht auf Prozessen, die auf mehreren Zeitebenen ablaufen - die Lernenden müssen den Referenten eines Sprechers im Moment identifizieren, eine Zuordnung zwischen dem Label und dem Referenten kodieren, sich an mehrere Labeling-Ereignisse erinnern und diese integrieren und ihre gelernten Zuordnungen nutzen, um das Objekt in neuen Kontexten zu identifizieren (Yurovsky et al., 2017). In den letzten Jahrzehnten haben sich zahlreiche Forschungsstudien mit der Frage nach "Wort-Welt"-Zuordnungen und Mechanismen befasst, die für die Entwicklung der Referenz verantwortlich sind (Hollich et al., 2000). Bisher hat jedoch keine Studie untersucht, ob und wie die subjektive Gewissheit über die Wort-Referent-Beziehung den Lern- und Gedächtniskodierungsprozess beeinflusst.

Ziel dieser Studie ist es daher, das bestehende Wissen über das Lernen von Wörtern zu erweitern, indem untersucht wird, ob und wie die referentielle Gewissheit Enkodierungsprozesse beim Lernen von Wörtern beeinflussen kann. Insbesondere ist untersucht, ob die Gewissheit der Teilnehmer über die Bedeutung eines neuen Wortes einen Einfluss auf die Enkodierungstiefe hat.

Um diese Frage zu untersuchen, haben wir kurze Filme (d.h. Objekt-Benennungsepisoden) aufgenommen, in denen eine Person jedes Mal ein neues Wort in Gegenwart von zwei unbekannten Objekten nannte. Die Gewissheit der Versuchspersonen über die Bedeutung der neuen Wörter wurde durch die Variation von zwei Faktoren moduliert: 1) Objektidentität (durch die Verwendung von zwei identischen oder zwei verschiedenen unbekannten Objekten) und 2) Verwendung der referentiellen Hinweise (indem der Sprecher während der Benennung auf eines der Objekte zeigte oder nicht). Nach der Teilnahme an den Lernsitzungen ist die Kodierung der neuen Wörter bewertet.

Unsere grundlegende Hypothese ist, dass unterschiedliche Grade der Gewissheit über die Bedeutung des neuen Wortes die Tiefe der Kodierung beeinflussen werden. Wir gehen insbesondere davon aus, dass die Kodierung umso besser ist, je geringer die Unsicherheit über die Bedeutung des neuen Wortes ist. Auf der Grundlage dieser Annahme erwarten wir, dass die Kodierung am besten ist, wenn die beiden unterschiedlichen Objekte präsentiert werden, während die Person auf eines von ihnen zeigt, und möglicherweise auch, wenn zwei identische

Objekte ohne eine Zeigegeste gezeigt werden, weil es unter diesen Bedingungen einfach ist, den potenziellen Referenten zu identifizieren. Im Gegensatz dazu gehen wir davon aus, dass die Unsicherheit über die Wortbedeutung hoch ist, wenn die beiden identischen Objekte präsentiert werden und die Person auf eines von ihnen zeigt, oder wenn die beiden unterschiedlichen Objekte ohne Zeigegeste präsentiert werden, und folglich die Kodierung schlechter ist als in den vorherigen Bedingungen.

Wir haben vier Experimente durchgeführt, drei davon sowohl mit Online- als auch mit Präsenzversionen. Unsere Ergebnisse sind noch nicht ausreichend, um zu behaupten, dass die Gewissheit über die Bedeutung der eingehenden Informationen einen direkten Einfluss auf die Enkodierung neuer Wörter hat. Dennoch bieten unsere Ergebnisse einige interessante Erkenntnisse und einen aufschlussreichen Blick auf die allgemeine Beziehung zwischen Gedächtniskodierung und Gewissheit über die Bedeutung eingehender Informationen. Das entwickelte Paradigma könnte außerdem ein guter Ausgangspunkt für spätere Forschungen sein.

# Acknowledgments

I would like to express my deepest gratitude to Prof. Dr. Hanna Marno from Eötvös Loránd University in Budapest. Her unwavering patience, insightful guidance, and continuous encouragement have been instrumental throughout the years, far exceeding the responsibilities of a supervisor. I also extend my thanks to Robert Danyi from Central European University, whose invaluable assistance in conducting the empirical research greatly contributed to this work. It has been a true privilege to learn from and collaborate with both of you.

This work is dedicated to my father, Ivan, who was always my guiding light and would be the proudest today.

# Chapter I. Introduction

## Structure of the thesis

The introductory part (Chapter I) presents the primary motivations for researching this topic and the main goals of this study.

The theoretical part (Chapter II) briefly introduces the premises behind our empirical study. First, the most prominent views on the notion of reference within the philosophy of language are going to be discussed. Secondly, the most significant views and empirical studies within the word learning research are presented. Finally, the idea behind our empirical study is introduced.

The empirical part of this thesis (Chapter III) thoroughly explains the material we developed for this project, our paradigm, Study 1 with five experiments and Study 2 with two experiments we conducted using the designed material. Finally, statistical analysis is presented, and results are discussed.

In the concluding part (Chapter IV), the thesis will make the case that, based on our results, at the moment there are insufficient grounds to claim that certainty about the meaning of incoming information would have a direct impact on encoding novel words. Furthermore, paradigm limitations, weak points, remaining open questions, and further research directions are discussed.

# Interdisciplinary aspects and research motivation

Within this project, we will provide insights into the relationship between memory encoding and certainty about the meaning of incoming information by using an interdisciplinary approach. This thesis will look at the topic through the optics of relevant fields such as linguistics, philosophy of language, psycholinguistics, cognitive psychology, and psychology.

Understanding those factors that could influence the encoding of novel words at a very early stage, when identifying the meaning does not necessarily occur, will enrich our knowledge about general mechanisms involved in new word learning. Furthermore, we aim to tackle further questions about strategies employed in early word learning (i.e., first language acquisition). We believe that every factor contributing to word learning, every aspect of the input, and the learning context that could predict later language development is significant. Such findings could illuminate the nature of language learning strategies and mechanisms, inspire new scientific questions, improve language therapy and early education fields, and maybe even help us attenuate the consequences of learning disabilities.

# Chapter II. Theoretical background

## 1. Introduction: Why are human languages so special?

*Humboldt: "The people's speech is their spirit. And their spirit is their speech. "*

About a half-century ago, the American linguist Charles Hockett (Hockett & Hockett, C., 1960) pointed out that all human languages possess several unique and remarkable properties that are absent in the communicational systems of all other creatures on the planet. He called them design features of the language (Hockett & Hockett, C., 1960). First, all languages have words, and all languages have thousands of words. Second, all languages have devices for acquiring new words. Third, all languages have ways of modifying the meanings of words (for example, by changing the time tense). They all have negation. They can form questions, permit abstractions and hypothetical, counterfactual, conditional, unreal, and fictional utterances. Furthermore, they all exhibit open-endedness, the ability to produce new utterances without difficulty. Finally, all languages permit displacements, the ability to talk about things other than the here and now. Our words symbolize and stand for objects, actions, and events. They function utterly differently from sounds. Words have a "stands for" relation to what they label. They do not need to occur contemporaneously or spatially coterminous with the objects, actions, and events they represent. Instead, they stand for what they label, even in a context far removed from original usage. Many of these magnificent properties of human language are described by the notion of reference.

## 2. What is the reference?

Reference represents words' map to objects, actions, and attributes and could be defined as a relation between specific representational tokens and objects. Nevertheless, when we think about the seemingly simple process of labeling some person, object, or act - something that we humans, as speakers, do naturally every day, many questions arise. Let us take, for example, the name of the famous linguist Noam Chomsky. When we say "Noam Chomsky," do we refer to a renowned linguist, MIT professor, political activist, philosopher, somebody who will be often mentioned in this work, somebody whose lectures we listen to, or maybe even somebody who looks like him? On the other hand, when his family member says his name, the same referential term might mean a husband, brother, or father. So, to which degree (if at all) can reference depend on subjective and private features of the speaker, such as her emotions, intentions, or other mental states? Are referential terms just referents, or rather something more? What are their meanings? Furthermore, how does the linguistic reference relate to the act of referring? How do referential terms come to refer? Is there a single reference mechanism common to all referring terms?

This chapter will briefly overview different viewpoints on these often-debated issues within the philosophy of language. For both clarity and brevity, the author will not go too deeply into the detailed debates regarding the meaning and syntactic form. Instead, the focus will be on the most relevant topics for our empirical study: firstly, on an overview of some general viewpoints about the nature of linguistic reference, and secondly, on the reference mechanism in the word learning process.

## *Notion of reference*

The relationships between reference and meaning, language and reality, are some of the most vigorously debated problems in the philosophy of language. In The Stanford Encyclopedia of Philosophy (2022), Michaelson, Eliot, and Marga Reimer offered an interesting categorization of different opinions about the notion of reference. Namely, they categorized different ideas about the reference as 'positive' and 'negative' views by looking at how philosophers have attempted to account for the referential features of linguistic terms and the general perspective of different viewpoints (Michaelson et al., 2022). 'Negative' views consider a reference to a non-scientific, subjective notion, or they even claim that our words fail to refer to in any sense (Michaelson et al., 2022). 'Positive' views assume a reference to be a genuine, substantive relation (Michaelson et al., 2022). In other words, the notion of reference is defined as something existing and is considered worthy of philosophical scrutiny. In further text, when discussing different philosophical views on reference, we will lean on this categorization.

## *Negative views of reference*

Let us start with two arguably most famous reference sceptics, Williard Quine, and Noam Chomsky. Chomsky's and Quine's thoughts about reference are wildly different, yet they both represent well what we in this work call 'opposing' ideas of reference.

In his works (1960), Quine introduced today's well-known philosophical problem, claiming that the world offers infinite possible word-to-world mappings. His famous example involved a linguist in a foreign country, observing a native pointing to the rabbit and saying "gavagai" at precisely the moment the rabbit ran by. But what is the linguist to think that "gavagai" refers to? Does it refer to the rabbit as a whole, to the rabbit's legs, to the rabbit's hopping, or even to the marks the rabbit leaves on the ground? Quine argues that for any assumptions we might have about the speaker of a given language, there are many competing possibilities regarding what their words refer to. In other words, empirical theories about referential meaning are underdetermined by their supporting evidence. Further, he argues that there is no general way of adjudicating between all the world's possibilities. However, Quine does not go so far as to claim that our words fail to refer, only that it makes sense to speak of what our terms refer only relative to some purpose we might have in assigning referents (Michaelson et al., 2022).

Even if very influential in the philosophical debate, Quine's interpretation argument is irrelevant to Chomsky, who has systematically debated many features of it. In his work, Chomsky completely rejects the externalists' interpretation of reference. He says a referential relationship between words and the world is not scientific; it is a "kind of neo-scholastic picture" (Chomsky, 2012, p. 28). Chomsky (2000) devotes many reflections to the criticism of the notion of reference. Let us quote one: 'A good part of contemporary philosophy of language is concerned with analysing alleged relations between expressions and things, often exploring our intuitions about the technical notions "denote," "refer to," "true of," etc. said to hold between expressions and something else. But there can be no intuitions about these notions, just as there can be none about "angular velocity" or "protein." [...] It is unclear that the theory of natural language and its use involves relations of "denotation," "true of," etc., in anything like the sense of the technical theory of meaning.' (Chomsky, 2000, p.130).

His theory concerning reference consists of two hypotheses: the idea that the relation of reference is a syntactic and internal notion and that speakers, not words, refer to objects in the world. He claims that denotation and reference are determined by beliefs and desires which interact with the faculty of language and determine how speakers use the formal expressions generated by the grammar in their speech acts (see, i.e., Chomsky, 2000, p. 70). Furthermore, for him, the notion of reference is not scientific; it is a pseudo-scientific notion that should not enter naturalistic analysis and experiments. Interestingly, Chomsky finds the notion of reference adequate for scientific and artificial languages (see, i.e., Chomsky, 2000.). Why? He does not give a specific answer to this question. Still, following his arguments, one could assume that the answer could be that in these languages, the reference is rigidly established through a convention, and symbols of these languages denote mainly abstract entities.

## *Positive views of reference*

After the described arguments, one question arises: if the reference is entirely internal, individual, and intentional – how can we explain that through language, speakers can express their thoughts, command, ask, compose poetry; namely, they can communicate? Many philosophers acknowledge that the descriptive content in question might vary across different speakers or even across time for the same speaker. However, they still consider a notion of reference and referential mechanism as an existing, genuine relation. In The Stanford Encyclopedia of Philosophy (2022), Michaelson, Eliot, and Marga Reimer introduced four ways how referential terms might refer to particular parts of our reality. They primarily focus on 'singular' referential terms - terms that, supposing they refer, refer to specific entities and individuals as opposed to groups or properties (Michaelson et al., 2022). They wonderfully describe and categorize "positive" philosophical views of reference in four models: descriptivist, causal, character, and intentionalist model (Michaelson et al., 2022). Since the four models offer an excellent categorization of the most essential philosophical views on reference, they will be shortly described in the further text.

In the descriptivist model, represented by Gottlob Frege (1892) and Bertrand Russell (1911), words are associated with a specific descriptive content that identifies a particular object or individual as the referent (Michaelson et al., 2022). In her mind, the speaker associates this content with the name in question. This way, the descriptive content uniquely determines the referent for the word. Words are associated with a description that isolates a particular object as the referent. So, the referents for parts of our reality are seen in the model of descriptions. Different versions of this model come from different ways of associating particular descriptions with specific uses of words. Accordingly, the reference in question may vary from one speaker to the next (Michaelson et al., 2022).

The routes of the second, causal model, could be already found in John Stuart Mills 'Millianism' (1867). However, the modern version was gradually developed and presented by Ruth Barcan Marcus (1961), Peter Geach (1969), Keith Donnellan (1970), and Kripke (1972). According to the causal model, words refer to the virtue of being associated with chains of use leading back to the initial use of the referent. The referential process is described this way: firstly, a speaker gives a name to an object. The speaker makes the initial connection by uttering the name and perceiving the object simultaneously. Then, after giving the reference, the name is passed on from speaker to speaker through communicative exchanges. The name, i.e., referent, becomes the "tag," and the subsequent speaker borrows the reference from the speaker earlier in the chain. Speakers succeed in referring to something employing the name because their uses of the specific name are links in a communicational chain going back to the initial connection of the object with the specific name (Michaelson et al., 2022).

Furthermore, the character model, introduced by David Kaplan (1990) offers a different line of thinking. Namely, any number of names with the same phonological or orthographic form refers to a different person or entity. According to this viewpoint, by looking only at how the word is written and pronounced, one cannot know what the term refers to. The listener must also see the speaker's intention to talk about one or another individual and to be familiar with the term the speaker wants to use in the given context. So, on the character model, names are, first and foremost, a mental phenomenon highly dependent on the speaker's mental state (Michaelson et al., 2022). Linguistic names help us externalize the use of a mental name to refer to in our thoughts (Michaelson et al., 2022).

Finally, following the intentionalist line of thinking, the Gricean approach to reference was introduced by Keith Donnellan (1968) and Gail Stine (1978), and is categorized as intentionalist model by Michaelson et al., (2022). In the intentionalist model, words refer to the virtue of being used intentionally to refer to particular objects (Michaelson et al., 2022). Differently said, references are uttered as part of complex intentional acts that target specific individuals or entities. Part of these intentional acts are our gestures, which make objects salient in a context. A similar could also be said for the group; for example, a term like 'he' can refer to whoever counts as the most salient male in a specific context. The basic idea is that the speaker's intentions to get the listener to identify a particular object as the referent could be a key to the connection between a term and its reference. These intentions aim to get the listener to identify some object as the referent and suggest how she intends to act (Michaelson et al., 2022).

In our empirical study, we intended to elicit different levels of referential certainty. We assume that using or not using referential cues and gestures in our intentional acts during word learning sessions will make some objects more easily (i.e., with more certainty) selected and remembered as referents for our referential terms. Therefore, we lean on positive views of reference, especially on the intentionalist model.

Furthermore, since words participate in a complex social system, in further text, this thesis will reflect on the notion of reference within the word learning process. We will do so by describing a few studies concerning our most prominent word learners: babies and small children.

## *The notion of reference and the word learning*

McMurray and his colleagues (2012) described the process of learning a word in four steps: (1) we have to identify a speaker's referent at the moment, (2) encode a mapping between the label and referent, (3) recall multiple labeling events and integrate across them and (4) use their learned mappings to identify the object in novel contexts. But what does it mean to really know a word? How does the referential mechanism influence the process of word learning?

Conventionally, 'knowing a word' requires knowing the sound form of the word (its sequence of consonants and vowels) and its denotation, including its semantic reference and its syntactic properties (Swingley, 2009). Many famous studies in language acquisition clearly showed an early bias in perceptual categorization (Eimas et al., 1971) and strong adaptation to the native language phonology (Kuhl et al., 1992). Very early, infants can learn even the word forms that have no apparent relevance in the conversational interaction (Swingley, 2009). This remarkable ability is nicely illustrated in one of the classical experiments, where Jusczyk and Hohne (1997) tested infants' ability to learn word forms that have no apparent relevance in conversational interaction. Researchers visited 8-month-olds in their homes ten times over two weeks. At each visit, infants were played a 30-min recording of a woman reading three storybooks, and over the ten visits, children heard the stories read by five different women. Two weeks after the last visit, children were brought into the laboratory and tested for their preferences for the most frequent of the content words in the stories (such as jungle, python, and sneeze) versus foils with a similar English frequency of occurrence, and similar overall phonological shape, to the test words (e.g., camel, lanterns, and sloth). Children showed a significant preference for the words from the stories. How do children succeed in this demanding task? In answering this question, some researchers pinpoint the importance of linguistic cues (such for example intonational markers, clause boundaries, and consonantal patterns) and some extent of statistical patterns recognized in the speech they hear (such as, for example, frequency of common lexical groupings as the indicators) (Swingley, 2009). Either way, infants learn the sound forms of some words and imply different processes to do that (Swingley, 2009).
Nevertheless, at this point, another question arises: knowing the word would mean not only that we know what a word sounds like but also what the word means. Words are more than associations - words are references, often referring to abstract concepts that could not possibly be present in the naming contexts. Nonetheless, already between one and two years of age, children can refer to the abstract concept or category and make numerous generic expressions (for example, "Cats are cute" and "Cats have ears").

Moreover, many very early acquired words have no concrete, real-world referents (for example, "why," "hello," "never"). How could these words be associated only based on perceptual similarities between entities? Considerably fewer research studies have addressed whether infants understand the word they remember and recognize, as well as which ways they imply in the task of finding the reference. In the further text, some of them will be mentioned.

Preissler and Carey (2004) conducted an empirical study that nicely describes the distinction between an association and a reference. Experimenters introduced 18- and 24-month-old infants to a photograph of a novel entity and named it (using an unknown label, "whisk"). Afterward, they asked infants to extend the word to another picture of a whisk or an actual 3D whisk. On an associative account, one could expect that infants would select the photograph, as it is perceptually more similar to the picture that co-occurred with the introduction of the novel word. However, children rarely, almost never, selected only the photograph. Instead, the small participants chose either the object alone (approximately 50% of cases) or both the picture and the object (about 50% of cases). These results suggest that they understood something profound: words refer to concepts and are not tethered to their associated perceptual impressions (Waxman & Gelman, 2009).

Furthermore, it should be remembered that words participate in a complex linguistic and social system. With the attempt to establish a word's meaning, children consider many social cues, starting with simple ones such as the pointing or eye gaze and coming to the trustworthiness and intentionality of the speaker. Brooker et al. (2013) perfectly illustrated this effect in a study investigating whether the speaker's reliability affects children’s selective trust and learning. In this study, 18-month-old infants were exposed to a speaker who accurately or inaccurately labeled familiar objects. Subsequently, the speaker administered a series of tasks in which infants had an opportunity to: learn a novel word, imitate the speaker's "irrational" actions, and help the speaker obtain an out-of-reach object. In contrast to infants in the accurate (reliable) condition, those in the inaccurate (unreliable) condition performed more poorly on a word-learning task and were less likely to imitate (Brooker et al., 2013). Results showed that infants as young as 18 months could detect a speaker's verbal inaccuracy and use this information to attenuate their word recognition and learning of novel actions (Brooker et al., 2013). Furthermore, since each word is embedded in an incredibly complex linguistic system of human languages, in the aim of interpreting the meaning of the novel word, infants attend to its surrounding linguistic context (Waxman & Gelman, 2009), including the presence or absence of devices such as determiners (for example, in some languages, use of articles - "This is the dog" and "This is a dog"), and the number, variety, and relation among arguments (for example, "push" vs. "break").

In our empirical study, we aimed to elicit different referential certainty levels by varying factors in the surrounding linguistic context. Before going into details of our empirical study, this thesis will shortly reflect on the most relevant viewpoints concerning word learning and go through some prominent word-learning research studies in the next chapter.

# 3. Birth of the Word

Adults swim effortlessly through a sea of words, recognizing and producing tens of thousands daily. Children are immersed in these waters from birth, gaining expertise in navigating language over their first years. Before children can tie their shoes, thousands of words are already mastered (Hollich et al., 2000). Children's vocabulary rapidly expands by approximately 19 months of age, and they are suddenly transformed into master word learners acquiring up to nine words per day (Hollich et al., 2000). But what does it take to learn a word? How do children break the language barrier from first utterances to productive vocabulary?

Identifying the speaker's referent at the moment, finding the map between the label and referent, and recalling previous labeling events requires children to reason synthetically, putting together their emerging language understanding with their knowledge about both the world and the people in it. Many factors contribute to word learning, ranging from social information about speakers' intentions to biases that lead children to extend categories appropriately (Roy et al., 2015). However, are the attentional or social factors the leading ones in the word learning process? Do children have certain biases that constrain the hypothesis space and assist them in linking words to objects, actions, and events, or are only their general cognitive abilities - attentional and logic mechanisms - enough to make the necessary associations?

Responses to the questions mentioned above have polarized the field of language development. The most prominent viewpoints concerning word learning will be further described as an introduction to our empirical study. Contributions of generative approaches, social pragmatics, and associations' views will be described shortly. Furthermore, this thesis will go through many relevant word-learning research studies, primarily done in first language acquisition, again, with our most proficient language learners: babies and children.

## *Generative approaches*

One of the most remarkable characteristics of knowing and using a language is that speakers can both produce and comprehend an unlimited number of sentences with different syntactic structures and word combinations, some of which they have never encountered before. This striking property of human language, sometimes defined as its discrete infinity, was the starting point of generative linguistics (Eisenbeiss, 2009). Noam Chomsky (1957, 1959, 1965) argued that only the study of the linguistic system in the mind/brain of individual speakers could lead to an explanation of the fact that we can generate infinitely many different expressions using a finite number of stored elements. Proponents of generative approaches to language learning argue that the syntactic knowledge language learners acquire is undetermined by the input. Therefore, they assume an innate language acquisition device that constrains children's hypothesis space when they acquire their native language (Eisenbeiss, 2009). Although scholars often disagree on how general or domain-specific this acquisition mechanism is and whether it is fully available from the onset of language learning (Eisenbeiss, 2009), the assumption of initial biases that constrain children's hypothesis space was a central concept. This assumption also became the guiding principle behind many prominent theories of word learning. Some of them are the Constrains or Principle theories. In the further text will be presented some of these theories, which were the standpoints of our paradigm.

In naturally occurring everyday conversations, children must assign meaning to the new words they hear. As both philosophers (e.g., Quine, 1960) and psychologists (e.g., Clark, 1991; Golinkoff et al., 1994) have stressed, determining reference is a complex problem. When we hear a new word, how do we know if it refers to an object, attribute, action, or event? The shared explanation is that there are some strategies which narrow the infinite search space. Family of constraint or principal theories claim that first language learners are biased to make certain assumptions about the meaning of the word, and these pre-given biases help enable them to direct their attention to those aspects of the environment that need to be selected and attended to (Hollich et al., 2000).

Clark and Markman were among the first researchers who introduced the idea of word-learning principles or constraints into the literature. Markman (1991) argued that when children make random word-to-world mappings, they could easily attach a name to the object part, object substance, or the action in which the object is engaged. Still, the line of experiments conducted by Markaman and his colleagues showed that they tend to first attach the new label to the whole object rather than the object part or attribute (Markman, 1987, 1988, 1991). In the experiments, Markman and Wachtel (1988) presented 3-year-old children with novel names for objects with a salient part. The experimenter then labeled the object and later tested whether the children attached the label to the salient part or the whole object. If the whole object did not already have a name, children consistently labeled the whole object rather than the part. If, in contrast, the whole object had a name that children knew, these children were more likely to assign the label to the object part. These results are explained by Markman's Whole Object Principle and Mutual Exclusivity Principle (Markman, 1987, 1988, 1991). Under these principles, children assume that a word labels a whole object rather than an object part, attribute or "non-object" substance, and they attach novel labels they hear to the nameless whole object referent, avoiding having two names for things (Markman, 1987, 1988, 1991). Second suggestion was later firmly established in the literature and is often called Novel Mapping Strategy. Both adults and children in different research studies consistently named unfamiliar objects with a new label (rather than a familiar object), even when there were no explicit or pragmatic cues from the speaker (Merriman et al., 1995) and even without explicit feedback from the speaker (Golinkoff et al., 1992). Around the same time, Eve V Clark (1987) proposed the Principle of Contrast, stating that any difference in the form of language marks a difference in meaning, and the Principle of Conventionality, claiming that for specific meanings, there is a conventional form that speakers expect to be used in the language community. Clark identifies three key pieces of evidence for conventionality: children's tendency to use adult word forms, their awareness when others deviate from these forms, and their expectation that others use the same terms for the same objects or ideas. For example, even if a child has their own invented or simplified word for something, they often replace it with the adult form once they learn it. They also ask for the names of things from adults. and often spontaneously repair their own word choices after making the mistake (it is not a "woof-woof" but a "dog"). Furthermore, they even often notice when others deviate from standard language, expecting consistent terminology from people around them and tending to correct unconventional word choices. This tendency indicates a strong preference for conventional terms that align with the speech of people around them (Clark, 1987). Clark's observations provide a valuable framework for understanding how children learn language systematically, showing an innate drive for linguistic conformity and conventionality. Further, Waxman and Kosowski (1990) proposed Noun- Category Bias, suggesting that children favour category relations when interpreting the meaning of novel nouns. Certain constraints have also been proposed for verb meanings and the mapping between semantic and syntactic representations, claiming the existence of constraints that enable children to use syntactic information to acquire verb meanings (Eisenbeiss, 2009).

The proliferation of principles posited for early word learning inspired many research studies and a proposition of developmental models that tried to unite sometimes very polarized views in the literature. Golinkoff et al. (1994) proposed a developmental model of word learning organized into two tiers, each representing developmental shifts in how children acquire language. This model suggests that word-learning principles evolve as children grow, shaped by linguistic experiences and innate biases. The first tier comprises three foundational principles essential for early word learning: reference, extendibility, and object scope. These principles dominate during infancy and with children younger than 18 months, helping young learners to understand that words stand for objects, actions, and events (reference), that words can apply to multiple similar objects (extendibility), and that words typically refer to whole objects rather than their parts (object scope).

Numerous experimental studies employing young participants support these early principles. For instance, in line with the Principle of Extendibility, very young children, already around 18 months of age, extend word meanings to similar objects based on shared attributes, such as shape, colour, or smell (Hollich et al., 2000). The Principle of Object Scope, consistent with Markman's research, clarifies that words typically refer to objects rather than actions or events and to whole objects rather than their parts. The Principle of Reference, fundamental to Golinkoff's framework, states that the word symbolizes or stands for objects, actions, and events. According to this principle, words serve a unique symbolic function, distinct from other sounds associated with objects (such as a telephone's ring). This principle emphasizes that words carry symbolic meanings and is developmentally first notion of reference, acquired already before the first birthday (Golinkoff et al.,1994).

Golinkoff et al. (1994) argue that these three foundational principles—reference, extendibility, and object scope—are essential cognitive-perceptual abilities that enable infants to begin the word-learning process. However, during early development, these principles only allow children to learn words in a gradual, one-at-a-time manner.As children approach the vocabulary spurt around age two, they begin to employ second-tier principles that support more complex word learning and faster vocabulary expansion. The second tier includes conventionality (the understanding that words have culturally agreed-upon meanings), novel name–nameless category (the ability to assign new words to previously unnamed objects), and categorical scope (grouping words into broader categories). The Principle of Conventionality underscores children's understanding that words are used in a standardized way by their linguistic community. The Novel Name–Nameless Category principle, also known as the Novel Mapping Strategy, enables children to rapidly map new words to unfamiliar objects. For example, as already described in previous text, when presented with a familiar and an unfamiliar object, young children are likely to link a new word to the object they don't have a name for. The Principle of Categorical Scope further advances word learning by encouraging children to categorize objects not only by perceptual similarity but by taxonomic relationships. For instance, research by Gelman (1987) shows that by around three-and-a-half years of age, children can categorize a small black bird with a flamingo rather than a bat, recognizing both birds as part of the same taxonomic category despite their visual differences.

All principle or constraints theories conclude that the child is predisposed to make specific hypotheses over others about word meaning, thus reducing the problem of the indeterminacy of word-to-world mappings. Although there is consensus that children employ many of these strategies in a new word learning, the mechanisms that foster these developmental strategies are more controversial. Some explanations assume an innate basis, whereas others propose an experiential one (Lederberg et al., 2000). Furthermore, some authors claim that only after these initial assumptions are made the child, as a constructivist, develops into an expert word learner (Golinkoff et al. 1994; Hollich et al., 2000).

## *Social Pragmatic Theories*

Infants are situated in their social system from their first day of life, and there is no doubt that they learn their first words through social interaction, but the unanswered question is how much they use social information. Is social information the crucial component that guides the learning process?

According to most social-pragmatic theories, children leverage social information from the outset of word learning (Bloom et al., 1998). Pioneers of social pragmatics theories see them as skilled word learners actively participating in a structured social world, with cooperating adults as social partners limiting their hypothesis space. If the child can "read" the social situation and if the adult is attuned to the child, then word learning becomes a kind of apprenticeship in which the social environment "feeds" word-to-world mappings to the child in digestible portions (Hollich et al., 2000). In other words, according to these theories, too many options in word-to-world mapping do not plague children; they do not try and guess what the label refers to - adults are the ones who guess what the child is focused on and then provide an appropriate word (Hollich et al., 2000). Further, children are social beings who actively try to comprehend, communicate, and seek out information relevant to their interests, and adults are ready to provide it (Bloom et al., 1998).

There is a considerable amount of evidence that children can utilize social cues in the service of word learning. They follow the direction of the social gaze very early and are more likely to do so in the presence of other communicative signals (Senju et al., 2008). Furthermore, one of the rare modern longitudinal studies in this domain claims that individual differences in children's ability to follow gaze predict differences in vocabulary development (Brooks & Meltzoff, 2008). Studies in the theory of mind show that infants even appear to represent others' beliefs, which affect their expectations by 12 months of age (Kovács et al., 2010). A series of studies by Baldwin and Tomasello (1998) reported that 19-month-old children would not attach a label to an unidentified object if the speaker did not seem to be referring to the object. For instance, children did not attach the label even if the object was interesting if the object appeared at the same time the label was uttered, and if the speaker was touching the object, even if the object was interesting to them (Baldwin & Tomasello, 1998). On the other hand, interestingly, 19-month-olds would attach a label to an object even when the speaker refers to a (partly) hidden object (Akhtar & Tomasello., 1996). Furthermore, in their experiment, Baldwin et al. (1996) showed that 18 months old children could evaluate whether an adult uttering a label ("It is a toma!") intended to label the object to which the child was attending or some other object. Thus, children around 18 months of age appear sensitive to very subtle social cues when attaching the label to objects (Hollich et al., 2000). There is much evidence that children are skilled, socially sophisticated word learners more than willing to reach out to embrace the language and all the power it provides. Therefore, in our empirical research study and word learning sessions we have designed, we rely on using social cues, such as eye gaze and pointing.

## *The Associationist View*

Every day, infants see and hear speaker naming and interacting with a particular object, often in temporal synchrony. For example, during almost every meal preparation, the child-parent would probably name "bottle" while washing it or giving it to the child. These events happen multiple times daily, so one could argue that only attentional and logical mechanisms could be enough for the child to connect the word "bottle" to its reference. So, we can say that infants and children learn rapidly from exposure to language, in ways that are unique to humans, combining pattern detection and computational abilities (often called statistical learning) with special social skills (Kuhl, 2004.). Following similar line of reasoning, associonist -oriented researchers claim that perceptual saliency, frequency, and an infant's general cognitive mechanisms, like attention and association, combined with the social exposure to language may be the keys to successful word learning (Hollich et al., 2000). Therefore, these authors argue that any innate constraints may not be needed because there is no ambiguity in the word learning process, and this process is straightforward (Hollich et al., 2000). Children notice most salient objects, actions, and events in their environment. They associate the most frequently used label with the most salient candidate (Hollich et al., 2000). Children map their first words onto references only with the general cognitive mechanism. After, their cognitive mechanisms are also responsible for the new strategies that come with the experience and are used to support the more sophisticated later learning (Smith, 2000). Therefore, researchers with this standpoint claim associative mechanisms play a crucial role in early vocabulary development. Infants learn words only by mapping them to salient objects in their learning environment (Smith, 2000).

Many research studies support these views. McMurray and colleagues showed that the degree of temporal synchrony in object naming episodes predicts mapping for young infants (McMurray et al., 2012). Sloutsky and colleagues (2001, 2007, 2008) have introduced a series of experiments supporting this standpoint. Their main arguments were that cognitive processes do not depend on top-down conceptual knowledge early in development. The assumption is that language learning early in development is grounded only on powerful general learning mechanisms, primarily associative mechanisms (Slousky et al., 2001). Later in development, conceptual content becomes more important (Sloutsky et al., 2001). Their work rests on three core assumptions: that the only building blocks for words and concepts are sensory and perceptual experiences and that these experiences are operated strictly by general-purpose processes (including associative learning, similarity assessment, and attentional weighting). By investigating the contribution of linguistic labels and perceptual similarity to young children's induction and similarity judgment, they have introduced an exciting suggestion of a developmental shift occurring between 8 and 11 years of age when conceptual cues become essential for word learning (Sloutsky et al.,2001). In the first experimental study done by Sloutsky and his colleagues, 4- to 5-year-olds, 7- to 8-year-olds, and 11- to 12- year-olds were presented with triads of schematic faces (two test stimuli and target) which varied in perceptual similarity, with one of the test stimuli sharing a linguistic label with the target, and another having a different label. Participants were taught an unobservable biological property about the target and asked to generalize the property to one of the test stimuli. Although 4- to 5-year-olds' proportions of label-based inductive generalizations varied with the degree of perceptual similarity among the compared stimuli, 11- to 12-year-olds relied exclusively on labels, and 7- to 8-year-olds appeared to be a transitional group. Therefore, the author concluded that developmental shift happens between 8 and 11 years. It reflects in treating linguistic labels as an attribute contributing to the similarity to treating them as markers of a common category (Sloutsky et al.,2001). In further studies, they have investigated the suggestion of Noun- Category Bias by having children of 4 and 5 years of age participate in the two types of tasks -categorization and induction. Both tasks were performed with artificial animal-like categories in which appearance was pitted against category membership. Although the children readily acquired category-membership information and subsequently used it in categorization tasks, they ignored category membership during the induction task, relying instead on the appearance of items (Sloutsky et al., 2007).

These results support the idea that induction is only similarity-based early in development and that higher-level conceptual processes are unnecessary to account for the developmental evidence (Sloutsky et al., 2007). Moreover, the authors argue that the learning strategies available to infants are only a subset of those known to older children. There might be fundamental differences between how linguistic knowledge is acquired in infancy and later childhood (Swingley, 2009).

Recently, with technological advances, the scope of observational studies has broadened (Roy et al., 2006). It has often been argued that the ideal way to observe early child development is in the home, where the routines and context of everyday life are minimally disturbed (Roy et al., 2006), and the introduction of digital audio recording technology has enabled more natural, detailed, and unbiased observation. New powerful recording technologies coupled with novel data mining and modelling tools offer us a completely novel line of investigation and an opportunity to track influencing factors and the relationship between linguistic knowledge and a child's experience. With such comprehensive data, it is possible to precisely follow a child's subtle behaviours during word learning in the home environment. It is also possible to couple specific contexts with the word production, frequency of heard and produced words, and, finally, to build a model which can "step into the shoes" of a child and learn directly from the child's experience (Roy et al., 2006). Computational models of word learning came as a novel and exciting approach to nature-nurture discussion. Suggestions are that if a machine can be shown to acquire some capability or structure without corresponding innate preconditions, this provides evidence that the child's environment provides everything necessary and that only general cognitive mechanisms may be enough for the successful early language learning (Roy et al., 2006). Connectionists' models of word learning have provided an utterly novel methodology and line of investigation that has until now been unexplored. One of the first connectionist models made by Plunkett et al. 1992, mimicked natural vocabulary explosion even though the only processes involved were associative, without any a priori constraints or social skills built into this model (Hollich et al., 2000). Based on the functioning of their connectionist model, the authors argued that small and gradual changes in the underlying neural system could cause dramatic shifts like the vocabulary spurt (Hollich et al., 2000). Computational model CELL (Cross-Channel Early Lexical Learning) introduced by Roy and Pentland 2002 learned to segment and associate the spoken words with acquired visual shape categories based on transcribed speech and video input recorded during the infant-mother interaction. The model enabled quantitative analysis of the effects of visual context on speech segmentation and the impact of short-term memory size on word learning performance (Roy & Pentland, 2002). The most extensive effort to observe and computationally model the longitudinal course of language development of a child by now was introduced by Roy et al. 2005/2006. The Human Speechome Project offers more than 400 000 hours of data following the language development of a single child from birth to three years of age. The author's idea was to equip a child's home so that nearly everything the child hears and sees from birth to three years is recorded. Later, researchers had the idea to develop a computational model of language learning that takes the child's audio-visual experiential record as input. Finally, they aimed to evaluate the model's performance in matching the child's linguistic abilities as a means of assessing possible learning strategies used by children in natural contexts (Roy et al., 2006). The data emerging from these novel modelling perspectives persistently suggest that general cognitive mechanisms, primarily associative mechanisms, play a crucial role in early word learning. Association-oriented researchers capture the human infant's impressive capacity to attend to statistical regularities in their environments, highlighting children's powerful sensory and perceptual resources which they bring to the task of word learning. This work is valuable in many ways but needs to consider an impressive array of infants' conceptual capacities (Waxman & Gelman, 2009).

## *Premises on which our empirical study is based on*

Although there is a disagreement about how general or domain-specific different word learning mechanisms are and whether they are fully available from the onset of language acquisition (Eisenbeiss, 2009), there are some theoretical viewpoints which are widely and generally accepted in the literature. Therefore, these are the most critical premises we used as ground points in developing our paradigm:

1) Positive views of reference, especially the intentionalist model, stating that references are uttered as part of complex intentional acts and that an essential part of these intentional acts is our gestures, which make objects salient in a context.

2) Novel Mapping Strategy stating that the learners would refer novel words to a novel object (Merriman et al., 1995). In other words, they would always attach a novel label to the name-less object referent present, avoiding having two names for the same object.
3) Whole Object Principle, stating that children assume a word labels a whole object rather than the object part (Markman, 1991).
4) Mutual exclusivity principle, stating that children associate the novel label with the unfamiliar object, avoiding having two names for the same object (Markman, 1991).
5) Noun - Category bias, stating that children favor category relations when interpreting the meaning of novel nouns (Waxman & Kosowski, 1990).
6) Sensitivity to social cues, stating that humans can attend to social-referential information (eye gaze, pointing) in their environment from a very early age (Bloom & Markson, 1998; Golinkoff et al., 2000; Tomasello, 2000; Hollich et al., 2000) and that children use this social information in the service of word learning very early in the development (Tomasello, 2000; Golinkoff et al., 2000).
7) Associative mechanisms in the object naming situation, highlighting human's powerful sensory and perceptual resources brought to the task of word learning.

Based on these theoretical standpoints, we have developed the paradigm with the goal to evoke different levels of referential certainty in four different conditions. Namely, we expect that by varying object identity (i.e., either using two identical or two different unfamiliar objects) and by varying the use of the referential cues (having the speaker pointing at one of the objects or not), we are going to successfully modulate subjects' referential certainty (i.e., subjects' certainty about the meaning of the novel words). Our goal was to investigate if the referential certainty could be a factor that contributes to the encoding of novel labels (i.e., to word learning) at a very early stage, when encoding of the meaning does not necessarily occur.

# 4. Rationale of the study

Despite the progress made in understanding the cognitive, social, and linguistic components of word learning and factors underlying the development of the reference, many questions remain open.

Let us imagine a young child surrounded by all the world's exciting ventures. The child takes an afternoon walk with her mother in nature, and she is suddenly surrounded by the sounds of wind in the trees, birds singing, dogs barking, etc. Let us further imagine four hypothetical situations. The child sees a pigeon and a blackbird on the tree branch, looks at them standing next to each other, and finds them very interesting. While she has her busy hands on the stroller, her mother tells her: "Look, a pigeon!" Assuming that the child had not heard of this bird sort before and that both birds are unfamiliar to her, would the child assume that the name refers to the small black bird or the gray pigeon? If the child is unsure which bird the mother refers to, would this information be relevant enough to devote her attention and cognitive strength to remembering the new name? However, if the mother makes an effort, looks and points to the grey bird, and says, "Look, a pigeon," would this situation be more straightforward for the child and, therefore, more relevant in devoting her learning mechanisms to the task of making the connection and remembering the new label?

Let us imagine another possible everyday situation. The mother and her child saw two similar-looking pigeons on the ground. The birds are pecking seeds from the ground, which the child finds engaging. If a mother says, "Look, pigeon," relaxed while having her attention and hands-on pushing the stroller, would the child assume that the name refers to both (similar looking) birds? Would this assumption mean that the label "pigeon" refers to the category of grey similar-looking birds?

On the other hand, if the mother makes more effort and points only to one pigeon, would that be a confusing factor for the child? Would the child assume that, as her dog at home, the one bird could have a personal name, 'Pigeon`, or still guess that the category of similar animals is named so? No matter the final assumption, one could argue that only this tiny doubt in the first moment could influence early attentional mechanisms and make a difference in mapping between label and reference.

Undoubtedly, different mechanisms influence our word learning mechanisms very early when label mapping still does not happen. We already described some of these well-documented word-learning strategies in the prior text. For example, the Novel Mapping Strategy states that the learners would refer novel words to a novel object (e.g., Markman, 1991; Merriman et al., 1995 ), or similarly, Mutual Exclusivity Constraint suggests that the learners would assume that the objects do not have two names (Markman, 1991; Woodward & Markman, 1998), or, Whole Object Principle implies that word learners assume a word labels a whole object rather than the object part (Markman, 1991). Without the doubt, there are some word-learning strategies that help people make inferences in word learning situations. Still, research on how the certainty about the meaning of incoming information could influence learning processes remains relatively unexplored. Could the certainty about the meaning of incoming information be one of these early strategies which help us to limit infinite research space in making a connection between label and reference?

To answer this question, we have developed a paradigm to test whether the certainty about the meaning of incoming information would influence our attentional process and memory. Our paradigm employs recorded sessions of word-learning tasks in which two unfamiliar (i.e., abstract) objects are presented, and an actor offers the new label (unfamiliar word) in a labeling event. In order to modulate the subjects' certainty about the meaning of the novel words, we varied two factors:

1) object identity (by using two identical or two different unfamiliar objects)
2) use of the referential cues (with having the speaker pointing at one of the objects or not)

To evoke different levels of referential certainty, we have developed four main conditions.

In the first two conditions, two identical unfamiliar (i.e., abstract) objects are presented, and the referential certainty is modulated by having or not having the referential cue (i.e. pointing) during the object naming event. In the latter two conditions, two different unfamiliar (i.e., abstract) objects are presented, and the referential certainty is again modulated by having or not having the referential cue (i.e., pointing) during the object naming event.

Our project contains of two studies. Study 1 has five experiments, three of them tested in-person and two of them online. Study 2 two experiments, one tested- in person and one online. We have slightly varied the different components of the described learning task in four experimental trials under four conditions. Still, our goal throughout the study was to evoke different levels of referential certainty in different conditions to test the relationship between referential certainty and our memory in new word-learning situations. We hypothesized that different levels of certainty about the meaning of the novel word will influence the encoding depth (i.e., our memory for the novel words). Expressly, we assume that the lower the uncertainty is about the meaning of the novel word, the better the encoding will be. We expect that encoding will be the best when the two different objects are presented with the person pointing towards one of them and possibly when two identical objects are shown without a pointing gesture (conditions 1 and 2). In these conditions, identifying the potential referent will be easy. In contrast, when two identical objects are presented, and there is a pointing gesture towards one of them, or when two different objects are present without a pointing gesture, we assume that uncertainty about word meaning will be high (conditions 3 and 4). Consequently, encoding should be worse than in the first two conditions.

Our first goal throughout this project was to investigate whether referential certainty significantly influences our memory in learning new words. Our aim was to provide important insights about the relationship between memory encoding and certainty about the meaning of incoming information. We suggested that certainty about the meaning of incoming information is a significant factor that would impact novel word learning at a very early stage when encoding the meaning does not necessarily occur. Furthermore, we tried to enlighten the general relationship between referential certainty and learning (i.e., memory). In addition, this study will provide an opportunity to test the suitability of the paradigm to investigate similar research questions.

Although our participants in all four experimental trials were adults, with the paradigm we developed, we also hoped to tackle the question whether the influence of referential certainty on memory encoding is one of the biases present already during first language acquisition (i.e., early in development). In further text, our empirical study will be presented.

# Chapter III. Empirical Study

## 1. General Method

### *Paradigm*

The whole project consists of Study 1 and Study 2. Study 1 investigates influence of referential certainty on encoding in a learning situation where unfamiliar objects (i.e., abstract objects) are named with novel labels (i.e., pseudowords). Study 2, investigates influence of referential certainty on learning in learning situations where familiar, everyday objects are named with novel labels (i.e., pseudowords). Study 1 consists of five experiments, and Study 2 consist of two experiments. Within our project, we enclosed two experimental environments: a laboratory and an online setting. Two experiments of Study 1 (Experiment 2 and Experiment 5) and one experiment of Study 2 (Experiment 2) were tested in an online environment, and the other four experiments were tested in a laboratory setting of the Baby Lab at Central European University.

The first part (i.e., the learning phase) of each experiment consists of word-learning sessions, which the authors designed. Our idea was that by varying object identity and using pointing gestures, we would evoke different levels of referential certainty (i.e., different levels of confidence about the meaning of incoming information). With this goal, we have designed four experimental conditions:

1. Having two identical abstract objects, with pointing (i.e., referential gesture) to one of them, during an object naming event.
2. Having two identical abstract objects without pointing (i.e., referential gesture) during an object naming event.
3. Having two different objects, with pointing (i.e., referential gesture) to one of them, during an object naming event.

4. Having two different abstract objects without pointing (i.e., referential gesture) during an object naming event.

We have prepared and designed abstract, novel objects and unfamiliar, novel words (i.e., pseudowords) for the learning phase of our experiments in Study 1. Pseudowords were divided into three groups: familiar (presented in the learning sessions), distractor (different in one letter with familiar words), and novel (new words). In compliance with the four conditions, the movies of the word learning sessions were recorded using abstract objects and familiar word group of pseudowords. In the first learning experimental phase movies of object naming episodes were presented in random order. In Study 1 participants were instructed to focus on the novel objects they saw on the screen, and they always heard one pseudoword from the familiar word group in an object-naming event. In the Study 2, the participants were instructed to try to understand and remember pseudowords they hear in an object naming event. Test phase of all experiments in both Study 1 and Study 2 consists of a simple recognition memory task with a yes/no recognition test. Pseudowords (familiar, distractor, and novel words) were presented on the screen in random order, and participants had to answer the question of whether they had already heard the written word.

## *Hypotheses and goal setting*

Our main hypothesis is that encoding depends on referential certainty, i.e., that different levels of certainty about the meaning of the novel word can influence the depth of encoding. The assumption is that the least uncertainty is about the interpretation of the novel word, the better the encoding will be (see Fig.1).

In particular:

1) We expect that encoding will be good in the condition where two different objects are presented, and there is a pointing gesture towards one of them because uncertainty about the meaning of the novel word will be low (i.e., subjects will interpret the meaning of the word as a reference to the object at which it was pointed at).

2) In the condition where two identical objects are shown in the absence of a pointing gesture, uncertainty about the meaning of the word should be low again. Accordingly, we expect that

encoding will be good. (i.e., subjects will assume that the novel label refers to both objects, thus to the object category).

3) In the condition where two identical objects are presented and there is a pointing gesture towards one of them, our assumption is that uncertainty about word meaning will be high (leaving open several referential possibilities, e.g. that the word refers to something specific about that object, i.e. a proper name, or just a property of the target object, and not object category) and encoding of the novel word will be worse than in the first two conditions.

4) In the condition where there is no pointing gesture and two different objects are presented, referential uncertainty will be high again, since the novel word could refer to any of the two objects with the same probability, thus we expect that encoding will be again worse than in the first two conditions.

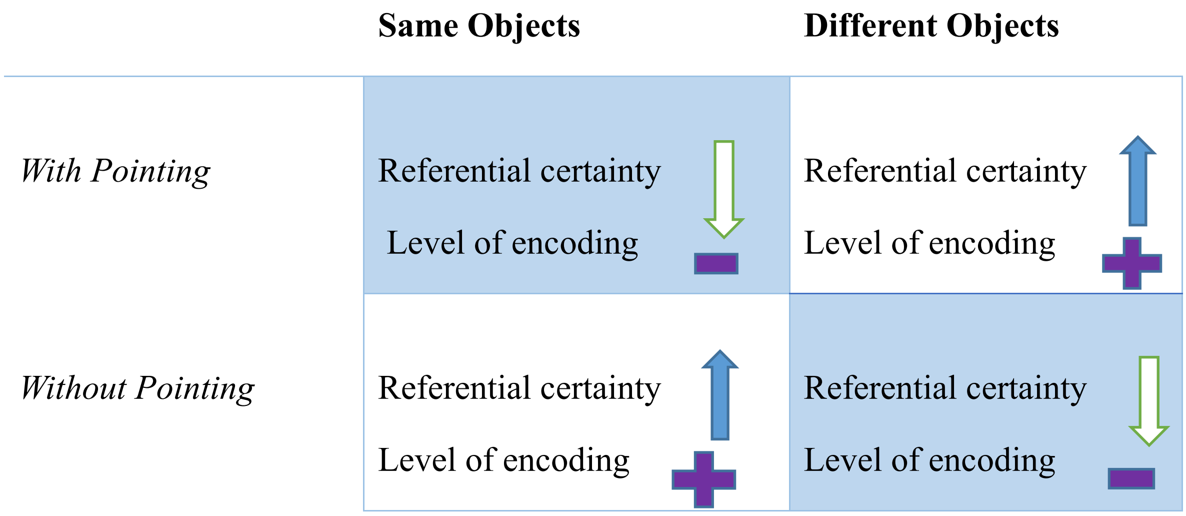



*Figure 1. Four conditions eliciting different levels of referential certainty about the meaning of incoming information*

## *Stimuli of the Study: Novel Objects*

The first step in designing our experiments and word-learning sessions was to make a database of three-dimensional, novel objects. Our goal was to make a database that could serve as stimuli for our experiments, and potentially be used in future studies, especially in the field of word learning studies. Our main criteria were:

1. Objects had to have similar visual complexity.
2. They were not supposed to be similar to any existing, familiar objects or living beings, for which subjects could already have a potential label.

In the following we will briefly explain the theoretical background of the preparation of our stimuli, and give a detailed description of the generated visual objects we used in our experiments.

## *Theoretical background of the preparation of the visual objects*

Biederman, in "Recognition-by-Components. A Theory of Human Image Understanding" (Biederman,1987.) argues that the ease with which we can code tens of thousands of objects is solved by mapping that input onto a modest number of primitives (36 components) and then using a representational system that can code and access free combinations of these primitives.

According to RBC's definition, the difference in the complexity of objects depends on the number of components required to look complete. In his experiments, Biederman varies the number of components required for some objects to look complete and then categorizes objects into groups with different complexity levels. Furthermore, he argues that both object complexity and the time necessary for object recognition primarily depend on the object's completeness. In other words, objects made of more components and recognized as complete objects require less time for identification than partial objects made of fewer components. Accordingly, by furnishing more combinations of components that could be simultaneously matched, complex objects can be more rapidly identified than simple objects. A complete and abstract object can be highly complex and asymmetrical, but the components are still simple volumes. If the components can be recovered and object perception is based on the components, then the whole object can be recognizable and remembered. Further, Biederman concludes that different arrangements of the same components can readily lead to different objects. Representational power derives from the enormous number of combinations arising from the modest number of primitives. Many research studies also show that humans rely heavily on shape similarity among objects for object categorization and identification (de Beeck et al., 2008.).

## *Preparation of the visual objects*

Following Biederman's theory about the primitive components of visual objects, we used the set of ten different types of objects, coming in ten different sizes, as fundamental elements for building abstract, unfamiliar objects. Accordingly, we could develop many combinations in building new, unfamiliar objects with only ten elements in ten different sizes (10X10 = 100 different elements, see Fig 2.). Using these elements, we first created 210 novel unfamiliar objects and, according to our exclusion criteria, reduced this set later to 192 objects.

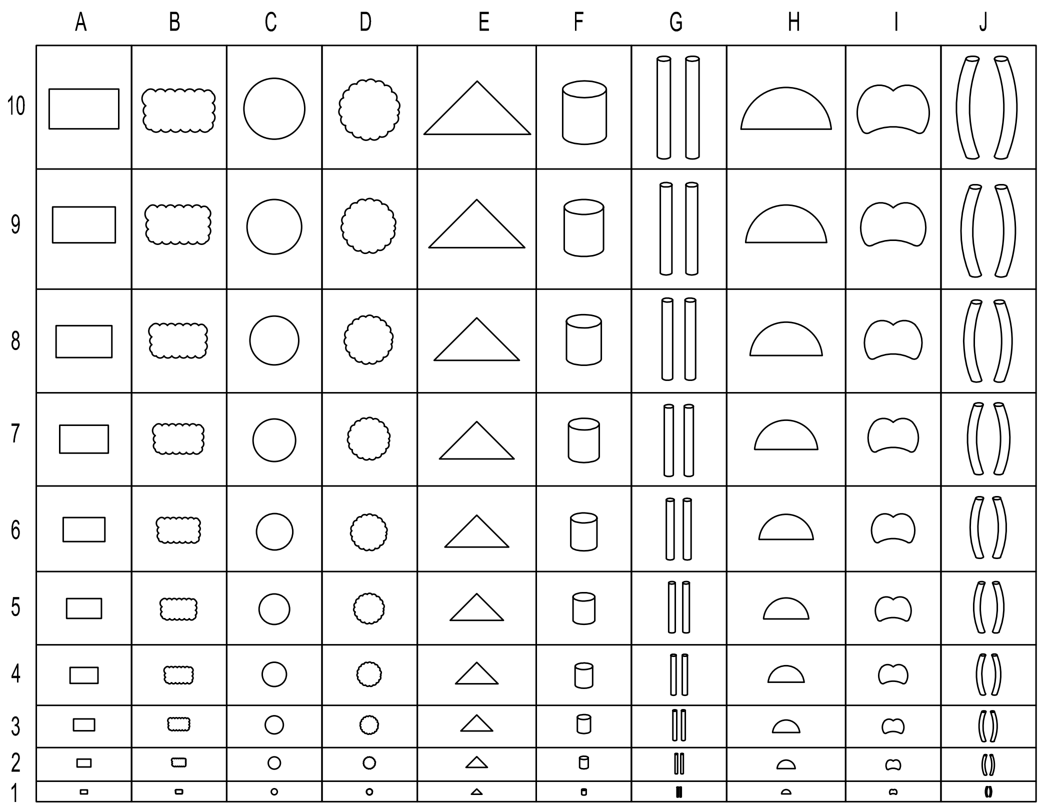
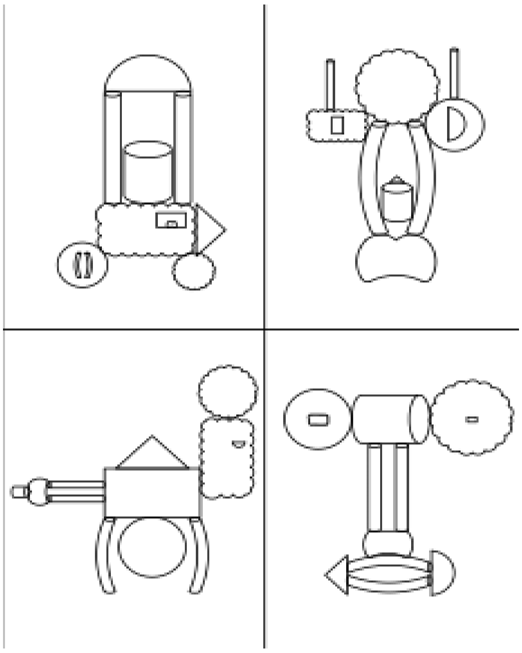



*Figure 2. Ten different elements in ten different sizes*

Firstly, we developed ten different size categories of elements (1,2,3,4,5,6,7,8,9,10, see Fig.2.). Elements in each category have the same surface area. Each size category is proportionally different from the next one for one scale factor. Furthermore, we developed ten different type categories of elements (A, B, C, D, E, F, G, H, I, J, see Fig.2.). Each type category has the same type of objects (exact symmetry, straight/curved edges, and shape). Our idea was that using one element from each size category (1,2,3,4,5,6,7,8,9,10) and each type category (A, B, C, D, E, F, G, H, I, J) for each new-unfamiliar object, each new object would have the same visual complexity, despite looking completely different (see some examples at Fig. 3.).



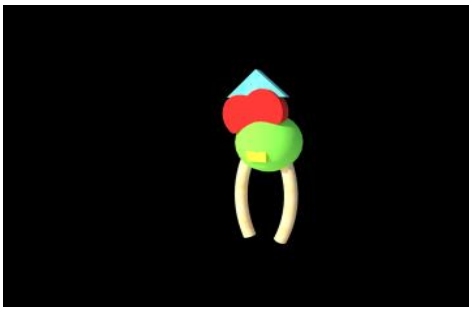


*Figure 4. Examples novel three-dimensional objects*

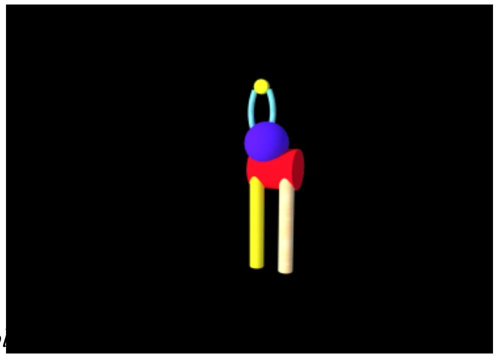


*Figure 5. Examp*

*3. Examples of novel two-dimensional objects*

Using SketchUp Pro drawing program, two-dimensional elements are coloured and transformed into three-dimensional elements (for example, see Fig. 4. and Fig. 5.). This set of hundred three-dimensional elements is used in creating 200 completely novel, unfamiliar three-dimensional objects. Our primary goals in creating our visual stimuli were that the novel objects should be visually salient, complete, and associative to some existing object categories in our memory to be memorable. However, at the same time, they were not supposed to be too similar to any existing everyday objects, so we could enable word mapping and learning new labels. Furthermore, to avoid similarity mixing mistakes, novel objects were supposed to be different from each other but simultaneously have similar visual complexity. Therefore, we tested our goals by conducting a small pilot study.

In the pilot study, five participants (two students and three Laboratory members of the Cognitive Science Department at Central European University) were asked to check our Novel-Object Database to exclude items that would be rated too similar to any existing object categories. If at least two pilot study participants found one object similar to the same existing object category, we excluded that object from our Novel-Object Database. This way, we could exclude 8 unfamiliar objects, getting the final set of 192 unfamiliar, novel, three-dimensional objects that were used in the project.

See the pictures of all 192 novel objects in the Supplementary Material.

## *Stimuli of the Study: Familiar, Distractor, and Novel pseudowords*

As novel labels for our abstract objects in word learning sessions, we have created 192 pseudowords. They are thought to be similarly complex, associative to existing words in the Hungarian language to be remembered, but still not too similar to any existing words. Our pseudowords were divided into three groups:

1. Familiar words are pseudowords presented during the short movies/learning sessions and which were supposed to be remembered during our word learning sessions (e.g., NACSU, MODI, LODI, BESU, CIBA).

2. Distractor words differ from familiar words in only one letter. Therefore, each familiar word has its distractor word pair (e.g., NACSU-DACSU, MODI-LODI, BESU-BETU, CIBA-CILA). They were not presented in our word learning sessions.

3. Novel words are novel and unfamiliar (i.e., unheard) pseudowords with similar complexity as the first two groups (e.g., JÖZI, LINÜ, NISA, TUGE, MORI, KAVÓ). They were not presented in our word learning sessions.

See the list of all the pseudowords we used in the Supplementary Material.

## *Stimuli of the Study: Videos of the Word learning sessions*

Firstly, we recorded videos in the controlled environment at Baby Lab, Central European University. The background was black, and the table with the black cover was placed in front of a female character in the videos. The person first smiled, waved, and then pointed to a black table on the left or right in both directions simultaneously or did not point at all, just continued to look into the camera. For our second recording, the person looked down after the smile and hand-wave before pointing to the left or right. Afterward, we recorded the audio of another female person (CEU Cognitive Science Department Laboratory member and native Hungarian language speaker) saying "Szia"- "Hello" (in Hungarian) and "Nezd"- "Look" (in Hungarian) and saying all our prepared familiar words. The recorded videos and audio were edited in the VSDC Free Video Editor. Using the video editing program, we have prepared 12s (for the third experiment) or 13s (for all other experiments in both studies) word learning sessions. We added prepared pictures of our three-dimensional objects, simple grey occluders to cover them in the first part of the video, and audio ("Szia, "Nezd," and one familiar word per video). Each video-word-learning session starts with a hand wave and the greeting "Szia"- "Hello" (see Fig.6.). Further, in each movie, a person says "Nezd" - "Look." At the start of the video, novel objects are hidden behind occluders (see Fig.6.). After the occluders move, the person provides a novel word in the presence of two unfamiliar objects. Object exposure time in each experiment in both Study 1 and Study 2 was 4s except the Experiment 3 (Study 1), in which exposure time was 3s.

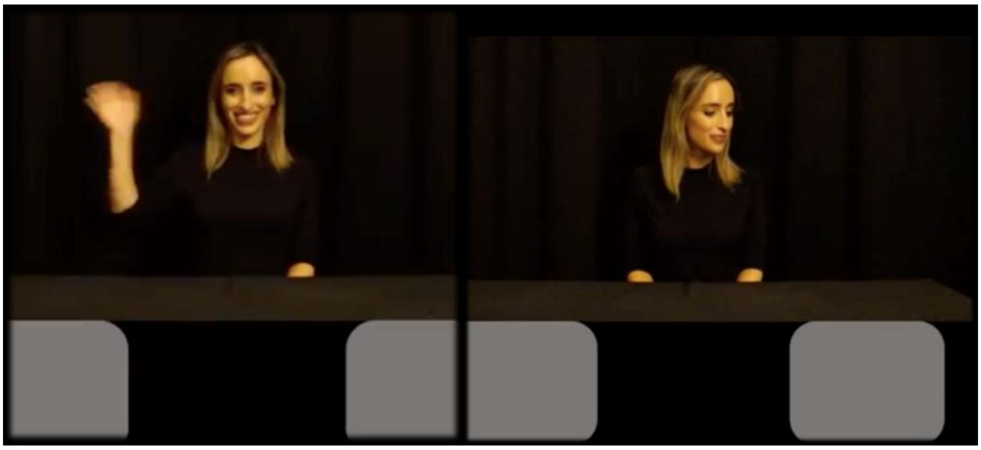


*Figure 6. Greeting and initial inspection*

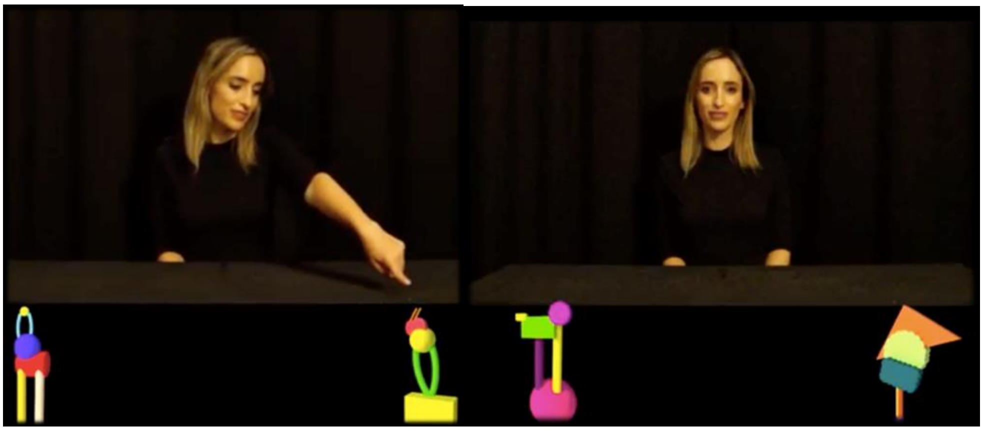


*Figure 7. Pointing and No Pointing Conditions*

# *Presentation of the Study*

We used the PsychoPy open-source package to run the Python experiments. All experiments had three main parts: a greeting of the subjects and a short instruction on the screen, stimulus presentation, and recognition test phase. First, participants were greeted and instructed to pay attention to the objects. Then, during the stimulus presentation phase, movies of the word-learning episodes were presented randomly on the screen. In Experiment 1 (Study 1), the stimulus presentation phase contained 64 randomly presented movies (in four different conditions, containing 64 different words). In all other experiments in both studies., the stimulus presentation phase contained 32 movies in four different conditions, containing 32 different words. During the recognition test phase, Familiar, Distractor, and Novel words were presented on the screen.

The recognition test phase of the first experiment (Experiment 1, Study 1) contained 192 words (64 familiar, 64 distractor, and 64 novel words), and all later experiments' recognition test phases contained 96 words (32 familiars, 32 distractor, and 32 novel words). After each word, participants were supposed to answer a question, "Did you already hear this word?" written on the screen, with "Yes" or "No" alternatives on the keyboard.

Because of the Covid 19 epidemy and the world events it caused, most of all, extended lock-downs in Hungary in 2020., we had to take a long break in our in-person laboratory testing at Central European University. These events motivated us to transform and adjust our experiments for online testing. Participants were recruited, and our experiments were presented on the Pavlovia.org and Testable Minds (Testable.org) platforms. Therefore, Experiments 2., and Experiment 4. of Study 1 were tested online, as well as Experiment 2 of Study 2.

## *General Procedure*

Participants were sitting in a quiet room, in front of a computer screen where the movies of the learning sessions and test phase were presented. Subjects were instructed to press different keys on the keyboard to indicate their answers. The entire procedure lasted around 20 minutes (depending on the subjects' speed).

During the learning (i.e., experimental) phase, movies of object naming episodes were presented in random order. Participants were instructed to focus on the novel objects they saw on the screen. During the test phase, distractor, novel, and familiar pseudowords were presented on the screen in random order. Participants had to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives. We collected data both about number of correct answers and subjects’ reaction times. The only exclusion criteria in all four experiments were participants' reaction time above 4s.

# 2. Study 1

## *2.1 Experiment 1*

### *Stimuli*

### *Word learning sessions*

Learning sessions of Experiment 1 contained four conditions, two with and two without referential cues:

1. The 'First Pointing condition' contained short waving first while looking into the camera, then looking down and pointing to the object on the right, simultaneously with the instruction "Look," i.e., "Nézd," and the object's name. Half of these videos had different abstract objects (we will refer to this condition as the Pointing/Different Objects condition, see Fig.8.), and half had the same abstract objects we will refer to this condition as the Pointing/Same Objects condition, see Fig.9.) presented. In each video, the person named different pseudowords from a familiar word group.
2. The 'Second Pointing condition': contained short waving first while looking into the camera, then looking down and pointing at the object on the left simultaneously with the instruction "Look", i.e., " Nézd," and the object's name. Half of these videos had different abstract objects (we will refer to this condition as the Pointing/Different Objects condition, see Fig.8.), and half had the same abstract objects (we will refer to this condition as the Pointing/Same Objects condition, see Fig.9.) presented. In each video, the person named different pseudowords from a familiar word group.
3. The 'No pointing conditions': contained short waving while looking into the camera and continuously looking at the camera without any other referential cues simultaneously with the instruction "Look," i.e., "Nezd," and the object's name. Half of these videos had different abstract objects (we will refer to this condition as the No pointing/Different Objects condition., see Fig.11.), and half had the same abstract objects presented (we will refer to this condition as

the No pointing/Same Objects, see Fig.10.). In each video, the person named different pseudowords from a familiar word group.

*Figure 8. Pointing / Different Objects condition*

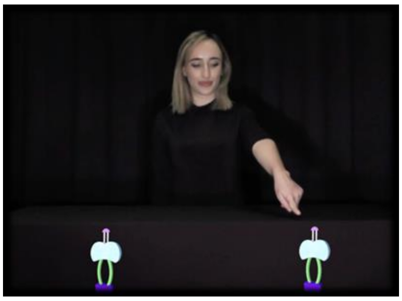


*Figure 9. Pointing /Same Objects condition*

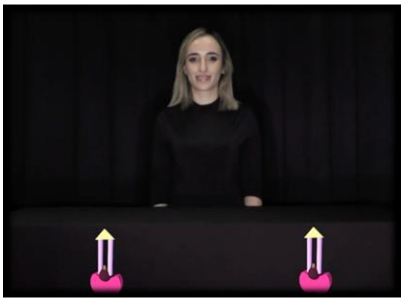


*Figure 10. No Pointing /Same Objects condition*

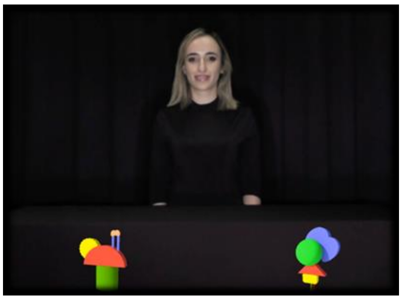


*Figure 11. No Pointing /Different Objects condition*

Learning sessions lasted 13s. Each learning session started with the person present but the objects hidden behind occluders on the screen. After the greeting and pointing gesture (or without the pointing gesture, dependent on the condition), the occluders were moved and two objects appeared. Again, objects were either the same or different depending on the condition. Object exposure time in each word-learning session was 4s.

## *Experimental Design*

Experiment 1 of Study 1 presented 64-word learning sessions and 64 novel pseudowords from the familiar-word group. Participants were instructed to focus on the novel objects; they were told that movies with novel/unfamiliar objects will be shown and that their task is to pay attention and try to remember them. Word-learning sessions in four different conditions were presented in random order. During the word recognition phase, distractor, novel, and familiar pseudowords were written on the screen in random order (192 words in total). Participants were instructed to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives.

## *Participants*

In Experiment 1, 35 male and female adults participated, between 18-60 years of age. All participants were native Hungarian language speakers. They were recruited via an advertisement published at Central European University and were rewarded for their time and participation with a small compensation. The criteria for inclusion were normal hearing, vision, and native knowledge of the Hungarian language.

## *Procedure*

Participants were sitting in a quiet room, in front of a computer screen where the movies of the learning sessions and test phase were presented. Subjects were instructed to press different keys on the keyboard to indicate their answers. The entire procedure lasted around 30 minutes (depending on the subjects' speed).

During the learning phase, movies of object naming episodes were presented in random order. Participants were instructed to focus on the novel objects they saw on the screen. During the test phase, distractor, novel, and familiar pseudowords were presented on the screen in random order. Participants had to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives. We collected data both about number of correct answers and subjects’ reaction times. The only exclusion criteria were participants' reaction times above 4s.

## *Data Analysis*

This study aimed to determine whether there are significant differences in the depth of the encoding/memory based on the correctness of the answers and depending on the different levels of referential certainty across four different conditions. It was hypothesized that the encoding would be better in conditions where two identical abstract objects were shown without a pointing gesture and where two different abstracts were shown with a pointing gesture toward one of them. In comparison, we expected that the encoding would be worse when two identical abstract objects were shown with a pointing gesture toward one of them and when two different objects were shown without a pointing gesture. We have collected data about the correctness of the answers for familiar, distractor, and novel words separately, and data about average reaction times. The collected data were analyzed using repeated measures ANOVA. We analysed two factors of object type - with two levels (same or different objects) and cue type with two levels (pointing or no pointing). Our exclusion criteria were participants' reaction time above 4s. No data were excluded in Experiment 1., both in online and in-person versions (see Fig 12.)

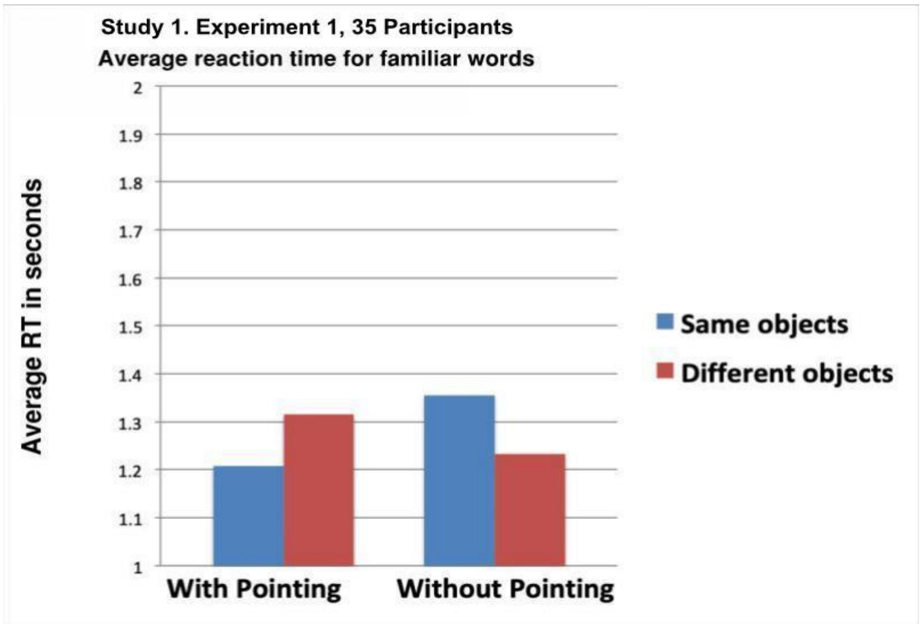



*Figure 12. Experiment 1: Average reaction times in seconds for familiar words*

# *Results*

The repeated measures ANOVA conducted for our in-person study of 35 participants (N=35) revealed a tendency for significant interaction between Object and Cue F(1,34) =1.977(p=.169) In the Experiment 1 , we did not find any main effect of Object type F(1,34)=1.5(p=.224) or a Cue F(1,34)=.102 (p=.751) for the familiar words. Similarly, we found no significant interaction between Object type and Cue F(1,34) =.000 (p=1.0) and no main effect of Object type F(1,34)=.012 (p=.913) or Cue F(1,34)=.094 (p=.761) groups across distractor word rejections. We have decided to run this t-tests because a tendency for interaction between object and cue was found. Bonferroni-corrected paired samples t-tests were performed to compare correct answers between the Same Objects/Pointing & Same Objects/No Pointing, Different Objects/Pointing & Different Objects/No Pointing, Same Objects/Pointing &Different Objects/Pointing, and Same Objects/No Pointing &Different Objects/No Pointing conditions, for both familiar and distractor words. Bonferroni-corrected paired samples t-test in for the familiar words revealed a significant difference between the Same Objects/Pointing and Different Objects/Pointing conditions (p = .044, 95% C.I. = -.1407, -.00212) (see Fig.13.). As predicted, the number of correct answers was higher in the Different Objects/Pointing condition (see Fig.13) than in the Same Object/Pointing condition or No Pointing conditions We found no significant differences in the performance for the familiar words between Same Objects/Pointing & Same Objects/No Pointing (p=.287, 95% C.I. = -14384, .04384 ), Different Objects/Pointing & Different Objects/No Pointing (p=.347, 95% C.I. =-.03642, .10071) or Same Objects/No Pointing & Different Objects/No Pointing (p=.799, 95% C.I. = -.07413, .09556. No significant differences between groups across distractor word rejections were found (See Fig.14). One sample t-test showed that correct novel word rejections significantly differed from the chance level t(34)=10.8, p<.000 (see Fig.15.).

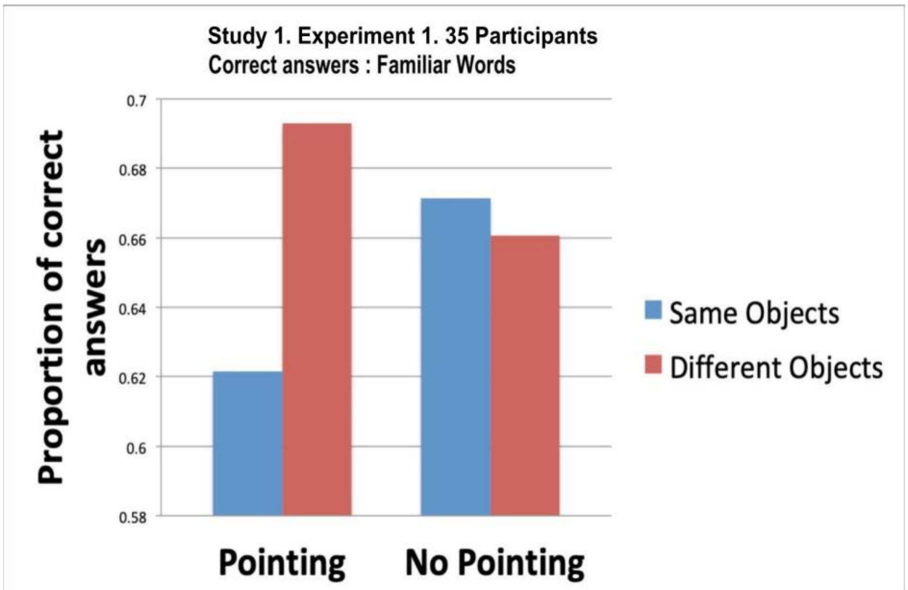



*Figure 13. Experiment I. Proportion of correct answers for the familiar words in the four conditions*

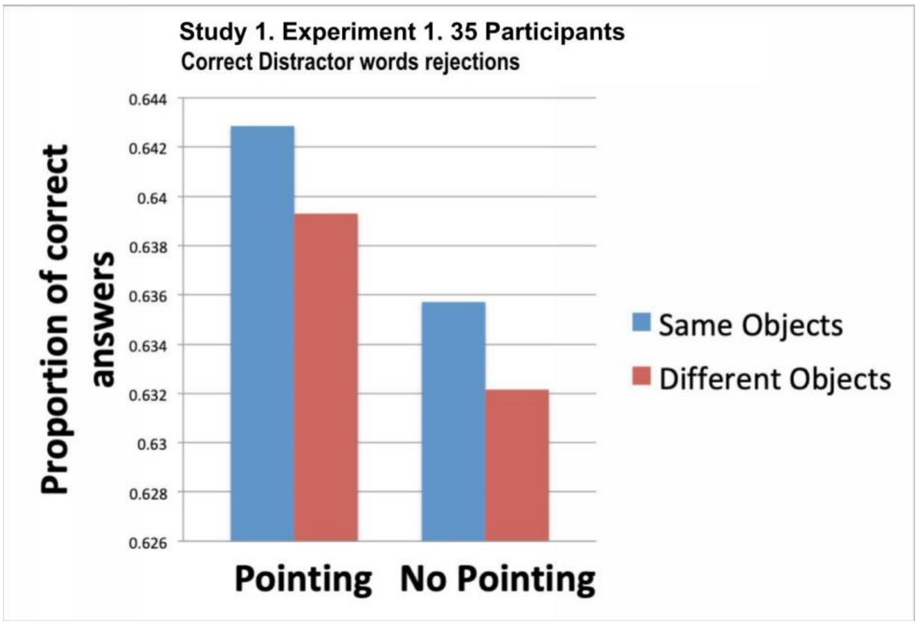



*Figure 14. Experiment 1: Proportion of correct rejections of the distractor words in the  four conditions*

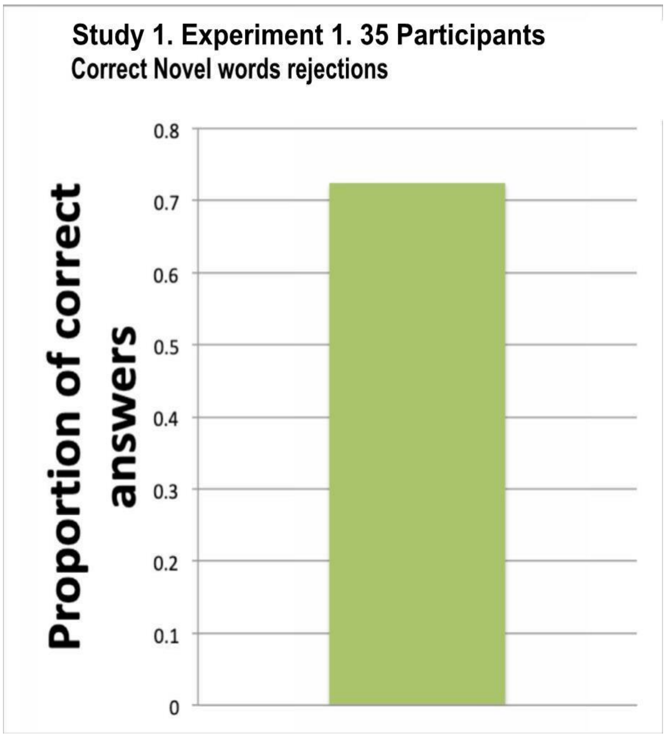



*Figure 15. Experiment 1,: Proportion of correct rejections of the novel words*

## *Discussion*

Experiment 1 aimed to investigate whether varying levels of referential certainty influenced the depth of encoding during novel word learning. We hypothesized that encoding would be more effective in conditions with higher referential certainty, specifically when two different abstract objects were shown with a pointing gesture and when two identical abstract objects were shown without a pointing gesture. Conversely, we expected poorer encoding in conditions with lower referential certainty, such as when two identical abstract objects were shown with a pointing gesture and when two different objects were shown without a pointing gesture.

Our repeated measures ANOVA analysis for the in-person study with 35 participants did not reveal significant main effects for object type or cue type for familiar words. However, a trend towards a significant interaction between object type and cue type was observed. This interaction suggests that there may be a subtle influence of referential certainty on encoding that did not reach statistical significance, at least in this sample size. Further analysis using Bonferroni-corrected paired samples t-tests indicated a significant difference in the proportions of correct answers between the "Same Objects/Pointing" and "Different Objects/Pointing" conditions for familiar words. This result supports our hypothesis, showing that higher levels of referential certainty, enhanced by different objects and pointing gesture to one of them, facilitates better encoding. However, no similar effect was found in a condition with two identical objects and no pointing to one of them, where we assumed that referantial certainty would be also higher.  Furthermore, no significant differences were found between other conditions , nor were there significant differences in distractor word rejections across conditions.The analysis of novel word rejections showed a significant difference from the chance level, indicating that participants could reliably reject novel words.

The findings partially support the hypothesis that higher referential certainty improves the encoding of novel words, but they clearly showed that conditions that contained some referential cues promoted better encoding. In other words, memory was not affected by referential certainty without the pointing gestures. The possible reason may be that the No pointing conditions in Experiment 1 did not include any referential cues, which could have led to failure in making a connection between the objects and the labels (i.e., pseudowords). Furthermore, after launching our first experiment and reflecting on our colleagues' feedback, we realized that the task with 64 learning sessions and 192 words (64 familiars, 64 novels, and 64 distractors) in the test phase may have been too demanding to the participants. In the following experiments, we aimed to make appropriate adjustments.

## *2.2 Experiment 2*

### *Stimuli*

Learning sessions were the same as in Experiment 1, but the number of presented movies was reduced to 32, in order to make the experiment easier for our subjects. Since Experiment 2 was launched in 2020, in the middle of Corona Virus Pandemic, we decided to adjust the experiment for online use. Learning sessions and the test phase were adjusted for online presentation using Pavlovia.org platform.

### *Experimental Design*

Word-learning sessions and word-recognition tasks were the same as in Experiment 1. However, as Experiment 2 was designed during the height of the COVID-19 lockdown in Hungary and much of Europe, our primary motivation was to make the necessary modifications to enable online implementation and testing. This online format allowed us to continue gathering data during the interruption of in-person testing. Furthermore, after running Experiment 1. and receiving feedback from the participants, we realized that our task in Experiment 1 (with 64 movies to watch and familiar words to remember) was too demanding to the participants, therefore we decided to modify our paradigm in order to make the task shorter and easier. Therefore, the number of presented word-learning sessions was reduced to 32. Accordingly, only 32 pseudowords of the familiar words were presented. Our new word recognition test phase contained 96 words written on the screen (32 familiar words, i.e., already heard, 32 distractor, and 32 novel words).

## *Participants*

In Experiment 2 23 male and female adults participated, between 18-60 years of age. The study was conducted in late 2020, during the lockdown in Budapest. All participants were native Hungarian language speakers. They were recruited via Pavlovia.org platform and were rewarded for their time and participation with a small compensation. The criteria for inclusion were normal hearing, vision, and native knowledge of the Hungarian language.

## *Procedure*

Experiment 2 was launched on the Pavlovia.org platform. Participants were recruited, signed the consent and took part in the experiment via the Pavlovia platform. The word learning phase, word recognition test phase, instruction for the participants, and task were the same as in Experiment 1.

## *Data Analysis*

Once again, we aimed to determine whether there are significant differences in the depth of the encoding/memory based on the correctness of the answers and depending on the different levels of referential certainty across four different conditions. It was hypothesized that the encoding would be better in conditions where two identical abstract objects were shown without a pointing gesture and where two different abstracts were shown with a pointing gesture toward one of them. In comparison, we expected that the encoding would be worse when two identical abstract objects were shown with a pointing gesture toward one of them and when two different objects were shown without a pointing gesture. We have collected data about performance in the word recognition task for familiar, distractor, and novel words (i.e., the correctness of the answers) and average reaction times for familiar, distractor, and novel words. The collected data were analysed using repeated measures ANOVA. We analysed two factors of object type - with two levels (same or different objects) and cue type with two levels (pointing or no pointing). Our exclusion criteria were participants' reaction time above 4s. No data were excluded in Experiment 2, tested online (see Fig.16).

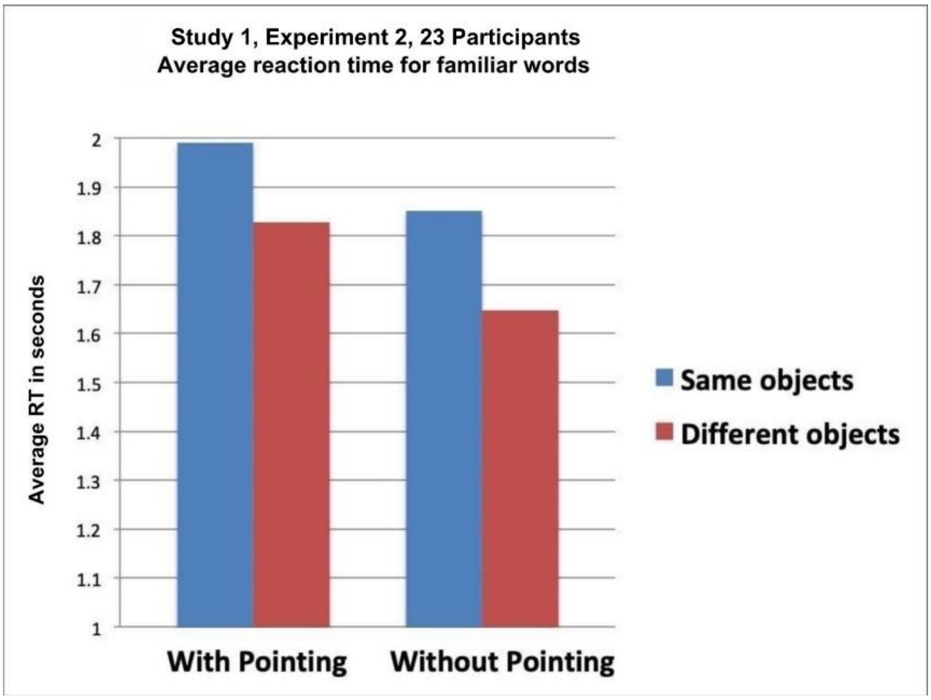



*Figure 16. Experiment 2 : Average reaction times in seconds for familiar words*

# *Results*

The repeated measures ANOVA conducted for our online study of 23 participants (N=23) showed no significant interactions between Object type and Cue F(1,22)=.075 (p=.787) for the familiar words. There were no significant interactions between the Object type and the Cue factors for the distractor words F(1,22)=.00 (p=1.00). No main effects of Object F(1,22)=.353 (p=.559) or Cue F(1,22)=.015 (p=.316) were found for the familiar words. However, we have found a marginally significant effect of the Object factor for the distractor words F(1,22)=4.2 (p=.053). Effect of the Cue factor for the distractor words was not significant F(1,22)=.083 (p=.777).Since Anova showed a marginally significant effect on the Object type, we have proceeded with the t-test. A Bonferroni-corrected paired samples t-test for the familiar words has not revealed any statistically significant differences between the Same Objects/Pointing & Same Objects/No Pointing (p=.638, 95% C.I. =-11628, .07281), Different Objects/Pointing & Different Objects/No Pointing (p=.318,95%C.I.=-.11524,.03915), Same Objects/Pointing & Different Objects/Pointing (p=.780, 95% C.I.=-.09057,.06883), or SameObjects/NoPointing &

DifferentObjects/NoPointing conditions (p=.580, 95% C.I. = -.12758, .07324) (see Fig.17.). We also found no significant differences between those groups for the distractor words (see Fig.18.). One sample t-test showed that correct novel word rejections significantly differed from the chance level t (22)=6.2, p<.000 (see Fig.19).

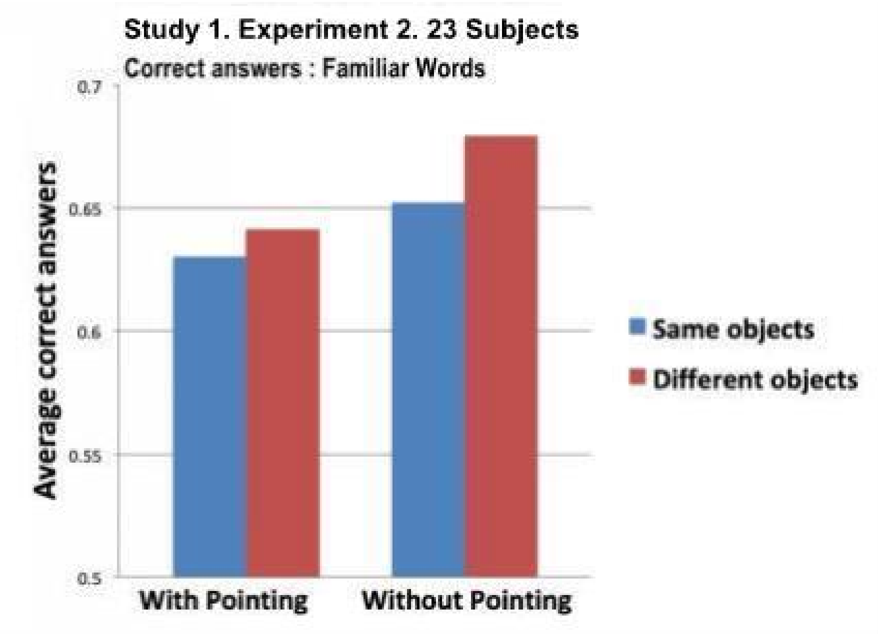



*Figure 17. Experiment 2: Proportion of correct answers for the familiar words in the four conditions*

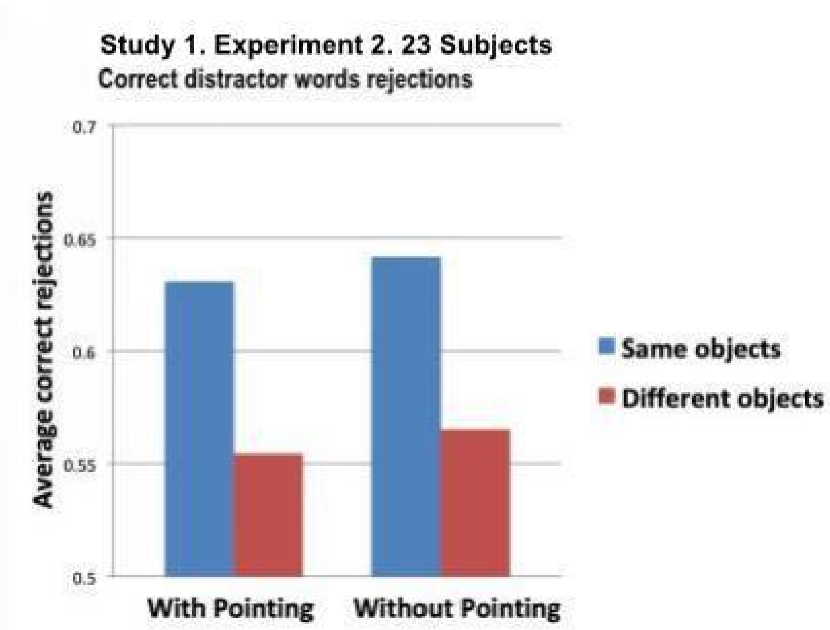



*Figure 18. Experiment 2: Proportion of correct rejections of the distractor words in the four conditions*

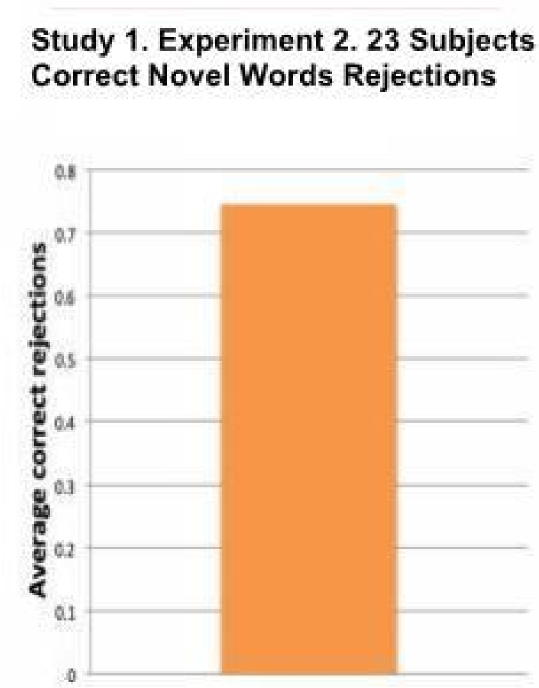



*Figure 19. Experiment 2: Proportion of correct rejections of the novel words*

## *Discussion*

The second experiment of Study 1, conducted online with 23 participants, aimed to assess the impact of referential certainty on the depth of encoding novel words. We hypothesized that higher referential certainty conditions—characterized by either two identical abstract objects shown without a pointing gesture or two different objects shown with a pointing gesture—would result in better encoding. Conversely, we expected poorer encoding in conditions with lower referential certainty.

The repeated measures ANOVA did not reveal significant interactions between object type and cue for familiar or distractor words. Additionally, no main effects were found for object type or cue in the familiar words group. However, a marginally significant effect of object type was observed for distractor words. Given the marginal significance in the object factor for distractor words, we proceeded with Bonferroni-corrected paired samples t-tests. These tests did not show any statistically significant differences between the various conditions in either the familiar or distractor words. However, a one-sample t-test indicated that correct novel word rejections significantly differed from chance level, suggesting that participants could reliably reject novel words.

## *Limitations of the first two Experiments and Subsequent Alterations*

The results from Experiment 1 provided partial support for our hypothesis. Specifically, memory for familiar words was better when two different objects were present with a pointing gesture towards one of them, compared to all other conditions. In other words, in conditions where pointing to one object accompanied the presentation of two different objects, higher referential certainty appeared to enhance encoding effectiveness. On the contrary, in the absence of pointing gestures (i.e., No pointing conditions) memory was not influenced by the object's identity.  In the absence of pointing gestures, no differences in encoding were observed regardless of whether the objects were identical or different, contradicting our expectations of improved performance with higher referential certainty in the Same Object/No Pointing condition. After launching the first experiment and considering the feedback we received, we concluded that the total lack of referential cues in the No Pointing conditions may have hindered participants' ability to link the labels with the objects, suggesting that participants may not have fully engaged with the labeling process in the absence of clear referential signals. Furthermore, we assumed that 64 learning sessions and 192 pseudowords within a test phase may have been too demanding and too challenging.

Experiment 2, conducted online, although the experiment was already simplified, still did not have any referential cues in No Pointing conditions. The Experiment 2 did not reveal any significant differences across conditions. These findings prompted adjustments in the following experiments which are elaborated upon in the subsequent sections of the text.

## 2.3 Experiment 3

### Stimuli

After launching the first experiments, we decided to make certain alterations in our stimuli for the Experiment 3:

Firstly, in order to elicit a stronger connection between the labelling events and abstract objects, we decided to have an initial inspection of the objects in each condition (See Fig 20). In order to do that, we have recorded new videos with slightly different scripts.

*Figure 20. Initial inspection of the objects behind occluders in each condition*

The new Pointing conditions with the initial inspection had the following scenarios:

1. Short waving first while looking into the camera, then looking down at the occluder (i.e., the object behind) on the left, then at the occluder (i.e., the object behind on the right, then looking back at the camera, then looking down and pointing at the object on the left.

2. Short waving first while looking into the camera, then looking down at the occluder (i.e., the object behind) on the right, then at the object (i.e., the object behind) on the left, then looking back at the camera, then looking down and pointing at the object on the left.

3. Short waving first while looking into the camera, then looking down at the occluder (i.e., the object behind) on the left, then at the occluder (i.e., the object behind) on the right, then looking back at the camera, then looking down and pointing at the object on the right.

4. Short waving first while looking into the camera, then looking down at the occluder (i.e., the object behind) on the right, then at the object (i.e., the object behind) on the left, then looking back at the camera, then looking down and pointing at the object on the left (See Figure 21. as an example).

*Figure 21. Example of the fourth new Pointing scenario*

The new No-Pointing conditions had the following scenarios:

1. Short waving first while looking into the camera, then looking down at the occluder (i.e., the object behind) on the left, then at the occluder (i.e., the object behind) on the right, then looking up into the camera (See Figure 22. as an example).

2. Short waving first while looking into the camera, then looking down at the occluder (i.e., the object behind) on the right, then at the occluder (i.e., the object behind) on the left, then looking up into the camera.

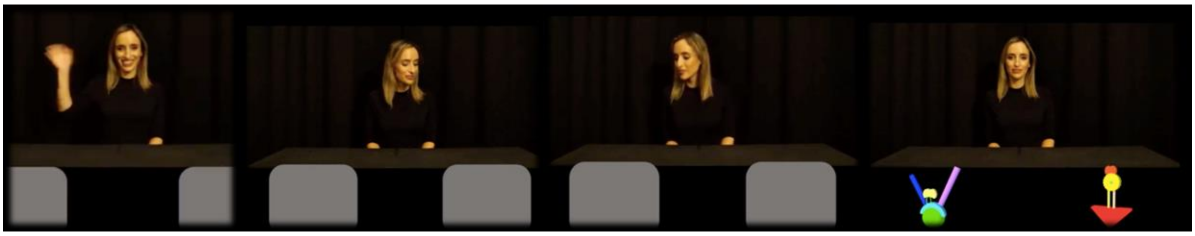


*Figure 22. Example of the first new No Pointing scenario*

We designed our new word-learning sessions to elicit a stronger referential connection between the labelling events and the abstract object or objects in each condition. However, the No Pointing conditions still lacked the later referential cue (pointing to one of the objects), leaving room for uncertainty about the novel label referent in certain conditions. After recording new videos, they were edited in the VSDC free video editor. Like for the previous experiments, occluders and abstract objects were added.

Secondly, we made some further changes with the videos of learning sessions. Occluders and abstract objects in learning sessions of Experiment 3 were 10% bigger in comparison to objects in learning sessions made for Experiment 1 and Experiment 2 (See Fig.23). Furthermore, the videos of learning sessions were 1s shorter than in all other experiments across two studies and lasted 12s. Objects exposure time was 3s (shorter than in all other experiments across two studies). Our goal with these changes was to make objects more noticeable (with the bigger size) and the whole experiment more simple and more engaging (by making the learning sessions shorter). In all other aspects, the movies of the word learning sessions were the same as in Experiment II. In each movie, the person in the video initially inspected both objects (the same or different, depending on the condition). Later, depending on the condition, during the naming event, the person pointed to the object or did not point to the object (same or different), with a goal to elicit different levels of certainty about the reference of the objects..

Finally, the number of presented word-learning sessions was again reduced to 32, similarly as for the Experiment 2. Accordingly, during the learning sessions, only 32 familiar pseudowords were presented. Our word recognition test phase contained 96 words written on the screen (32 familiar, i.e., already heard words, 32 Distractor, and 32 Novel words). Experiment 3 was tested in person at the Baby Lab, Central European University.

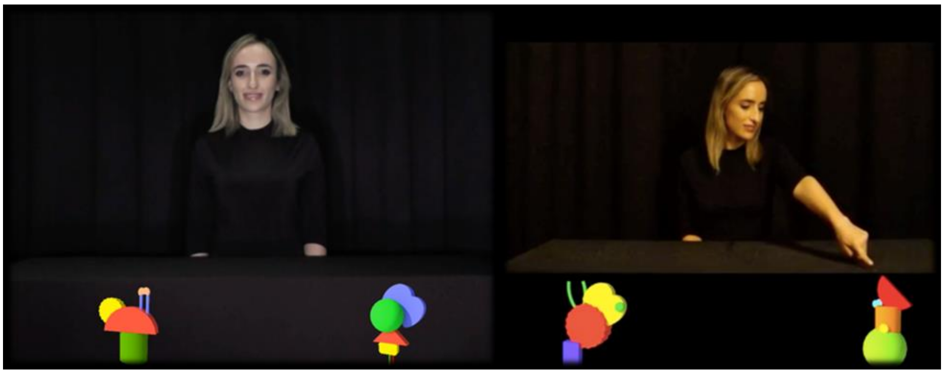


*Figure 23. Object size comparison between Experiment 1 - first picture and Experiment 3 - second picture*

## *Experimental Design*

Each experiment presented 32-word learning sessions and 32 novel pseudowords from the familiar-word group. Participants were instructed to focus on the novel objects; they were told that movies with novel/unfamiliar objects will be shown and that their task is to pay attention and try to remember them. Word-learning sessions in four different conditions were presented in random order. During the word recognition phase, distractor, novel, and familiar pseudowords were written on the screen in random order (96 words in total). Participants were instructed to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives. Except for the reduced number of word-learning sessions, the experimental design was the same as in In-Person version of Experiment I.

## *Participants*

In Experiment 3 participated 30 male and female adults between 18-60 years of age. All participants were native Hungarian language speakers. They were recruited via an advertisement published at Central European University and were rewarded for their time and participation with a small compensation. Like in other experiments, the criteria for inclusion were normal hearing, vision, and native knowledge of the Hungarian language.

## *General Procedure*

Participants were sitting in a quiet room, at the Central European University, in front of a computer screen where the movies of the learning sessions and test phase were presented. Subjects were instructed to press different keys on the keyboard to indicate their answers. The entire procedure with the changes we made for Experiment 3 lasted around 20 minutes (depending on the subjects' speed).

We collected data both about the number of correct answers and subjects' reaction times. Like in Experiment 1 and 2, the only exclusion criterion was participants' reaction time above 4s. No data were excluded in Experiment 3.

## *Data Analysis*

The goal of the third experiment was again to determine whether there are significant differences in the depth of the encoding/memory based on the correctness of the answers and depending on the different levels of referential certainty across four different conditions. We expected that encoding would be better in conditions where two identical abstract objects were shown without a pointing gesture and where two different abstracts were shown with a pointing gesture toward one of them. In comparison, we expected that encoding would be worse when two identical abstract objects were shown with a pointing gesture toward one of them and when two different objects were shown without a pointing gesture. The data about performance in the word recognition task for familiar, distractor, and novel words (i.e., the correctness of the answers) and average reaction times for familiar, distractor, and novel words were collected. The collected data were analysed using repeated measures ANOVA. We analysed two factors: Object type (with two levels: same or different objects) and Cue type (with two levels: pointing or no pointing).

## *Results*

The repeated measures ANOVA conducted for our Experiment 3 with thirty participants (N=30) revealed a marginally significant main effect of the Cue ($F(1,29) = 4.065$, $p = .053$) for the familiar words. Furthermore, the repeated measures ANOVA also revealed a marginally significant main effect of the Cue ($F(1,29) = 468.465$, $p = .094$) for the distractor words rejections. We have not found a main effect of the Object type either for the familiar words ($F(1,29) = .340$, $p = .564$) or for the distractor words rejections ($F(1,29) = .240$, $p = .628$). The analysis has not revealed significant interactions between the Object type and the cue neither in the familiar words group ($F\ (1,29) = .083$, $p = .776$) nor for distractor words rejections ($F(1,29) = .046$, $p = .832$). Because of the marginally significant effect revealed by ANOVA, we have continued the analysis with the t-test. Bonferroni-corrected paired samples t-test exhibited a marginally significant difference between the Same Objects / Pointing condition and the Same Objects/No pointing condition (p = .064, 95% C.I. = -0.00441, .14608) for the familiar words (See Fig. 24.). However, there was no statistically significant difference between the Different Objects/Pointing and Different Objects/No Pointing (*p* = .263, 95% C.I. = -.04288, .15121), Same Objects/Pointing and Different Objects / Pointing (*p* = .821, 95% C.I. = -.08283, .06616), or Same Objects/No Pointing and Different Objects/No Pointing (*p* = .579, 95% C.I. = -.11624, .06624 ) for the familiar words (See Fig.24). We conducted the same test to check the correct distractor word rejections level for the groups mentioned above and found no significant interactions between them (See Fig.25.). Furthermore, one sample t-test showed that novel word rejections significantly differed from the chance level (p = .000, 95% C.I. = .1941, .2809) (See Fig.26.).

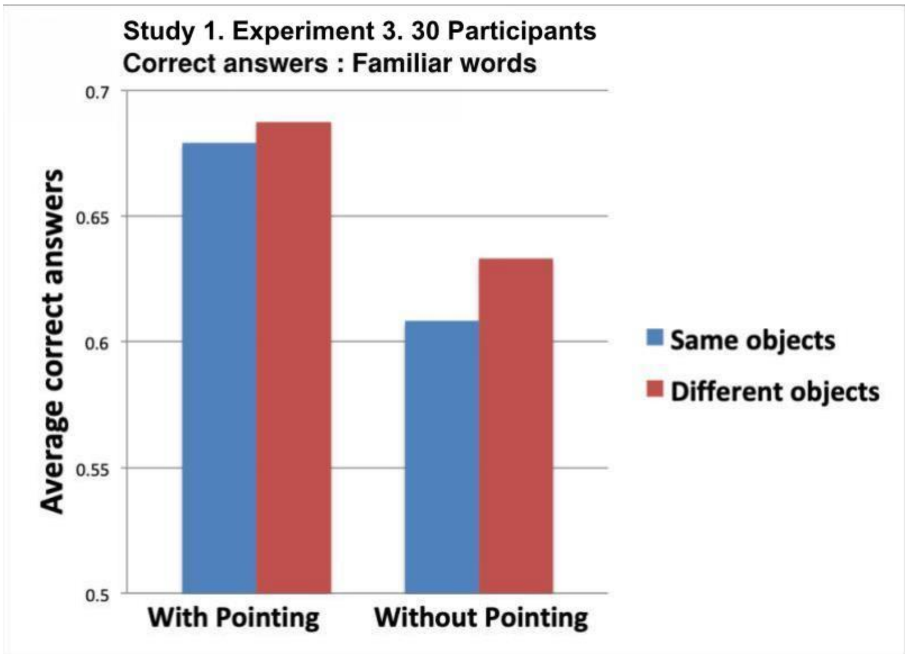



*Figure 24. Experiment 3: Proportion of correct answers for the familiar words in the four conditions*

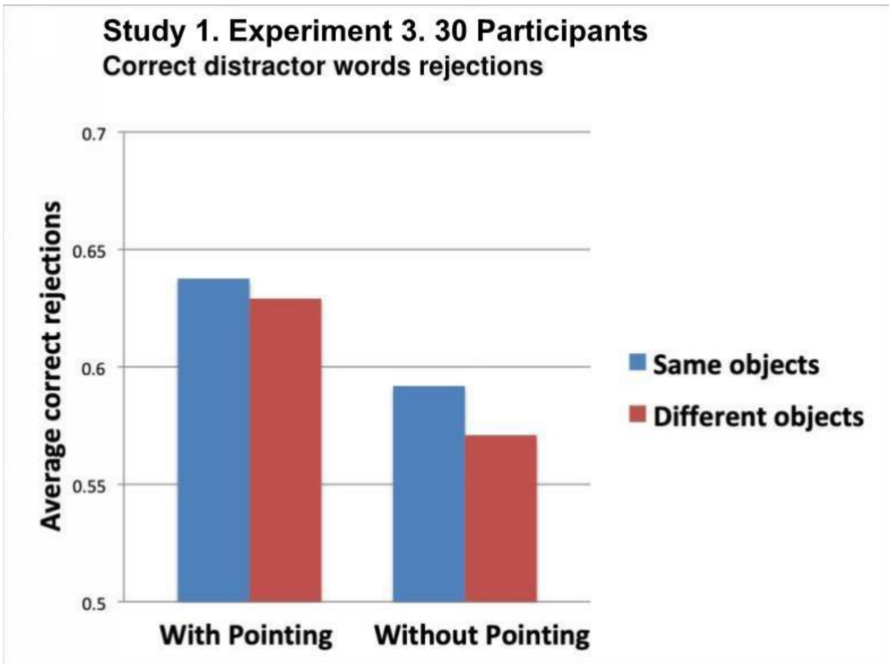



*Figure 25. Experiment 3: Proportion of correct rejections of the distractor words in the four conditions*

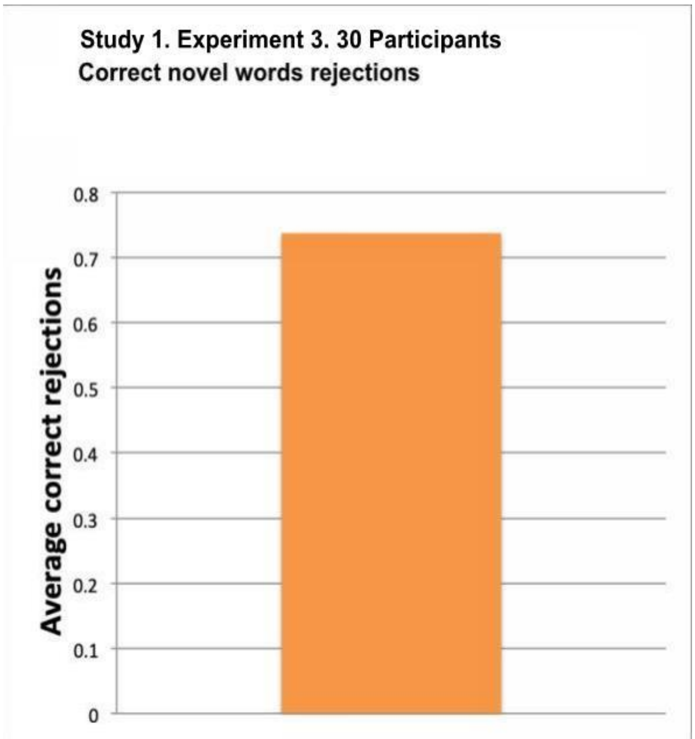



*Figure 26. Proportion of correct rejections of the novel word*

## *Discussion*

The goal of Experiment 3 was once again to determine whether different levels of referential certainty would significantly impact the depth of encoding/memory. Unlike the first two experiments, all conditions in Experiment 3 included a referential cue (initial inspection of the objects). Our hypothesis remained the same: we expected better encoding when two identical abstract objects were shown without a pointing gesture and when two different abstracts were shown with a pointing gesture toward one of them. Conversely, we expected worse encoding when two identical abstract objects were shown with a pointing gesture and when two different objects were shown without a pointing gesture.

The results of Experiment 3 revealed a marginally significant main effect of the cue, indicating that participants performed better in the Pointing conditions compared to No-Pointing conditions (See Fig.24.). This suggests that memory for novel words was better when the initial inspection was followed by a pointing gesture, which was true for both familiar words and the rejection of distractor words (See Fi.24. and Fig. 25.).

However, we did not find statistically significant differences in participants' performance based on the type of objects present in the No-Pointing conditions. Notably, participants performed better in the Same Objects/Pointing condition compared to the Same Objects/No Pointing condition, contrary to our hypothesis (See Fig.24.). This suggests that pointing towards one of two identical objects may have enhanced referential certainty and improved encoding in Experiment 3, even though this was not initially expected.

## 2.4 Experiment 4

## Stimuli

After launching Experiment 3, we decided to make additional alterations before starting Experiment 4. We have re-edited the videos and made changes which will be described in subsequent text.

Firstly, the size of the abstract objects and consequently, occluders, in the learning sessions was made to the original size, as it was in Experiment 1, Experiment 2, and Experiment 3 (See Fig. 23), because after launching Experiment 3 and receiving feedback, we found that the original size as more natural and less distracting. Secondly, videos were made to last 13s again, as was the case in Experiment 1 and Experiment 2.The duration of each learning session was increased again to 13s in total, hence, the object exposure time lasted again 4s in total.

In other aspects, the movies of the word learning sessions were the same as in Experiment 3. In each movie, the person in the video initially inspected both objects (the same or different, depending on the condition). Later, depending on the condition, during the naming event, the person pointed to the object or did not point to the object (same or different), potentially eliciting different levels of certainty about the reference of the objects.

## Experimental Design

In the learning phase 32-word learning sessions and 32 novel pseudowords from the familiar-word group were presented. Word-learning sessions in four different conditions were showed in random order. Like in the previous two experiments, participants were instructed to pay attention to novel/unfamiliar objects presented in the movies and that they should try to remember them. After watching the movies, the test phase followed. Distractor, novel, and familiar pseudowords were presented on the screen in random order (96 words in total). Participants were instructed to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives.

## *Participants*

In the Experiment 4, 45 male and female adults participated, and they were between 18-60 years of age. As in all other experiments, all participants were native Hungarian language speakers. They were recruited via an advertisement published at Central European University and were rewarded for their time and participation with a small compensation. The criteria for inclusion were normal hearing, vision, and native knowledge of the Hungarian language.

## *Procedure*

Participants were sitting in a quiet room at Baby Lab at Central European University. They were greeted, signed the consent form, and started with the experiment. Instructions, the movies of the learning sessions, and the test phase were presented on the computer screen. With the learning sessions lasting again 13s, the entire procedure duration was around 30 minutes (depending on the subjects' speed).

During the learning phase, movies of object naming episodes were presented in random order. Participants were instructed to focus on the novel objects they saw on the screen. During the test phase, distractor, novel, and familiar pseudowords were presented on the screen in random order. Participants had to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives, and they could indicate their answer by pressing different keys on the keyboard. We collected data about the number of correct answers and subjects' reaction times. The only exclusion criterion was participants' reaction time above 4s, and based on our criteria, no participants were excluded from the data analysis (See Fig.27).

## *Data Analysis*

The goal of Experiment 4 was again to determine whether there are significant differences in the depth of the encoding/memory based on the correctness of the answers and depending on the different levels of referential certainty across four different conditions. Our hypothesis was the same as in all previous experiments. The expectation was that the encoding would be better in conditions where two identical abstract objects were shown without a pointing gesture and where two different abstracts were shown with a pointing gesture toward one of them. In comparison, we expected the encoding to be worse when two identical abstract objects were shown with a pointing gesture toward one of them and when two different objects were shown without a pointing gesture. We have collected average reaction times for familiar (See Fig.27), distractor and novel words data about the correctness of the answers for familiar, distractor, and novel words. The collected data were analyzed using repeated measures ANOVA. We analyzed two factors of object type - with two levels (Same or Different Objects) and cue type with two levels (Pointing or No pointing). No participants were excluded from the analysis.

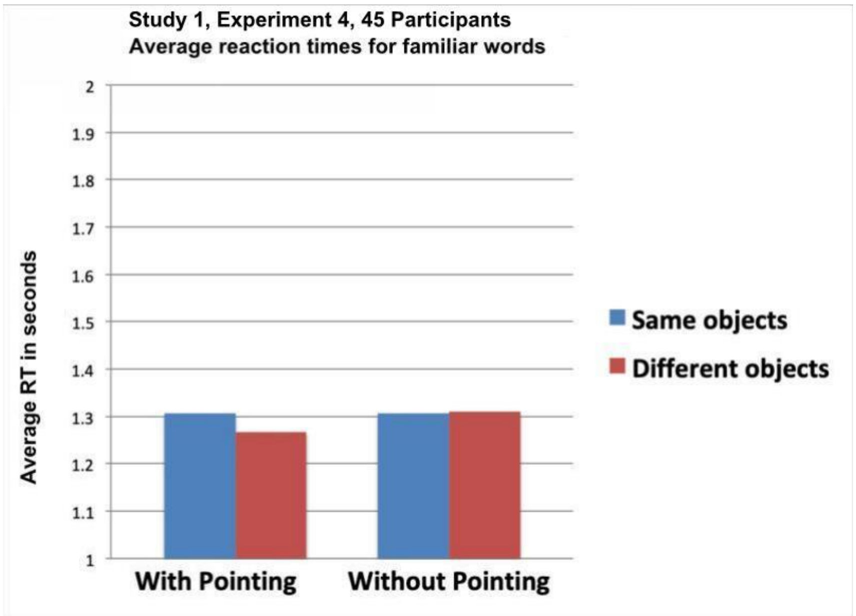



*Figure 27. Experiment 4: Average reaction times in seconds for familiar words*

# *Results*

The repeated measures ANOVA conducted for our in-person study of 45 participants (N=45) of Experiment 4 revealed a tendency for significant interaction between Object and Cue $F(1,44)=2.675$ ($p = .109$) for the familiar words. However, we have not found the main effect of the Object type $F(1,44)=.328$ ($p = .570$) or the Cue $F(1,44)=.176$ ($p = .676$) in the familiar words group. The repeated measures ANOVA for the distractor words showed a tendency for a significant effect of the Object type $F(1,44)=2.678$ ($p = .109$). Interaction between the Object type and the Cue $F(1,44)=2.021$ ($p = .162$) or the effect of the Cue $F(1,44)=.208$ ($p = .650$) were not statistically significant for the distractor words. Because of the marginally significant effect revealed by Anova, we have continued the analysis with the t-test. Bonferroni-corrected paired samples t-test, which we did for both familiar and distractor word, exhibited a marginally significant difference between the Same Objects / No Pointing condition and the Different Objects/No pointing condition ($p = .097$, 95% C.I. = -0.01001, .11556) for the familiar words (See Fig.28.). However, there was no statistically significant difference between the Different Objects/Pointing and Different Objects/No Pointing ($p = .173$, 95% C.I. = -.02268, .12268 ), Same Objects/Pointing and Different Objects / Pointing ($p = .497$, 95% C.I. = -.09864, .04864 ), or Same Objects/Pointing and Same Objects/No Pointing ($p = .432$, 95% C.I. = -.09844, .04288 ) for the familiar words (See Fig.28.). The same test we did for the correct distractor word rejections showed a significant difference between the Same Objects/Pointing and Different Objects/ Pointing conditions ($p = .045$, 95% C.I. = -.00221, .17002) (See Fig.29). We have not found any other significant differences between other groups mentioned above for the distractor words rejections (See Fig.29.). Furthermore, one sample t-test in our In-person version of Experiment III showed that novel word rejections significantly differed from the chance level ($p = .000$, 95% C.I. = .2034, .2799) (See Fig.30.).

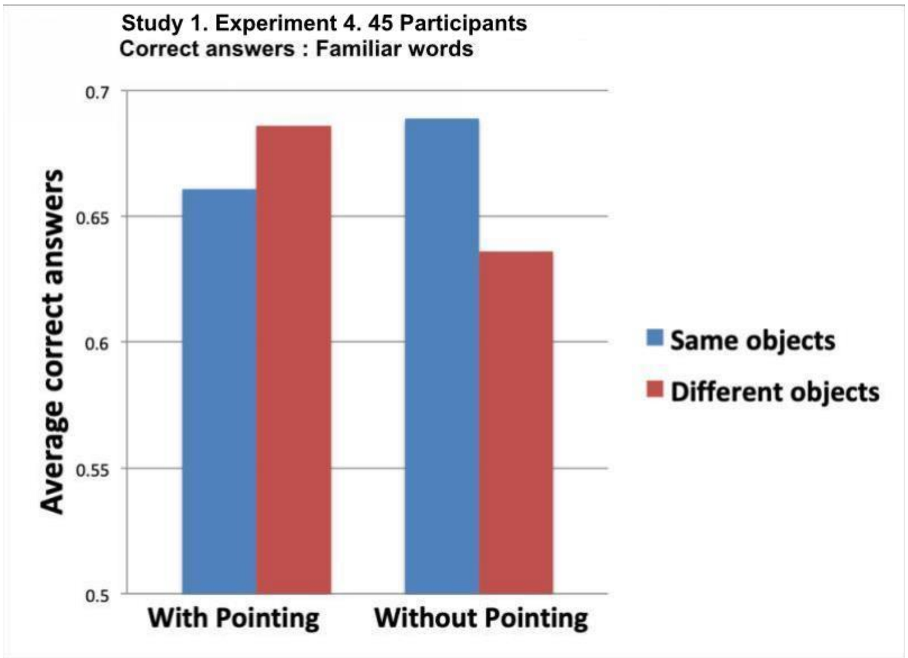



*Figure 28. Experiment 4 : Proportion of correct answers for the familiar words in the four conditions*

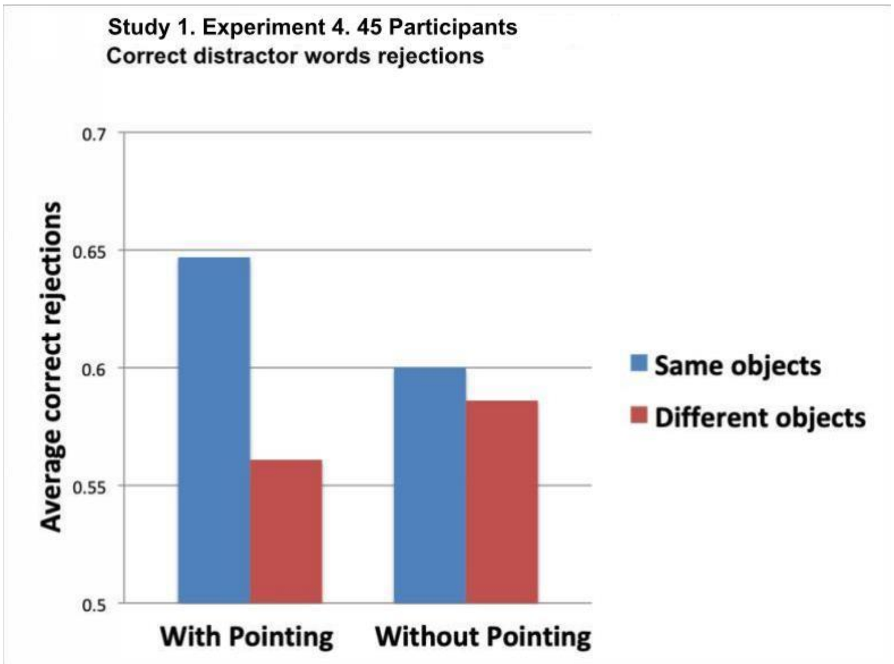



*Figure 29. Experiment 4: Proportion of correct rejections of the distractor words in the four conditions*

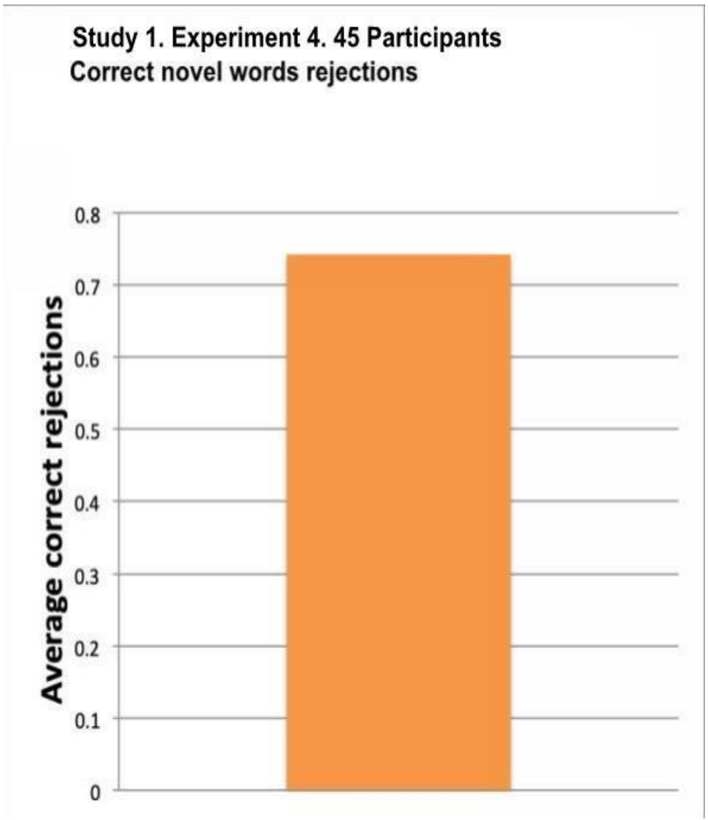



*Figure 30. Experiment 4: Proportion of correct rejections of the novel words*

# *Discussion*

The objective of Experiment 4 was to evaluate whether referential certainty affects the depth of encoding and memory performance across four conditions. The hypothesis remained consistent with previous experiments: better encoding was expected when two identical abstract objects were shown without a pointing gesture and when two different abstract objects were shown with a pointing gesture. Conversely, worse encoding was anticipated when two identical abstract objects were demonstrated with a pointing gesture and when two different objects were shown without a pointing gesture.

Results of Experiment 4 results showed marginal tendencies and some significant differences, but they did not entirely confirm our hypothesis. The interaction between object type and cue showed a tendency toward significance for familiar words, suggesting a possible influence of referential certainty. A significant difference in distractor word rejections between the Same Objects/Pointing and Different Objects/Pointing conditions indicates that pointing gestures may enhance the ability to reject distractor words when objects are identical, and therefore, the recognition was good in these conditions.

Overall, while some findings support our assumptions, the results were not conclusive. The marginal significance and specific conditions where effects were found only highlight the complexity of word-learning situations.

## *2.5 Experiment 5*

Due to the global COVID-19 pandemic and the strict government measures in Hungary, in early 2020, we were required to suspend in-person testing, leading to an almost two-year pause. To address these limitations and maintain data collection, we adapted Experiment 4 into an online format, creating Experiment 5. This online version allowed us to continue gathering data during the break in in-person testing, with the expectation that an expanded dataset from Experiment 5 would strengthen and potentially amplify the effects observed in Experiment 4. Experiment 5 will be described in the following section.

### *Stimuli*

Stimuli were the same as in the Experiment 4. However, stimuli were modified to be employed online and presented in the Testable Minds platform.

### *Experimental Design*

The experimental design of Experiment 5 was the same as in the Experiment 4. However, as in the other online experiments within Study 1, participants in the online study were working from home or a place of their choice, so we had no control over their working environment and working conditions.

## *Participants*

In Experiment 5, 49 male and female adults participated, and they were all between 18-60 years of age. The study was conducted in late 2020., during one of the Covid-19 lockdowns in Budapest. All participants were native Hungarian language speakers. Fifteen participants were students of Central European University. All participants were recruited via the Testable Minds platform (Testable.org) and were rewarded for their time and participation with a small compensation. As in the previous experiments, the criteria for inclusion were normal hearing, vision, and native knowledge of the Hungarian language.

## *General Procedure*

The Experiment 5 was practicly the same as Experiment 4, but was re-created and modified to an online study using the Testable Minds platform during the lockdown in Hungary in 2020. Participants were recruited, signed the consent and took part in the experiment via the Testable Minds platform. The learning sessions, word recognition test phase, instruction for the participants, and task were the same as in Experiment 4.

## *Data Analysis*

The goal of Experiment 5 was to assess the differences in encoding and memory based on the correctness of answers and varying levels of referential certainty across four conditions, mirroring the objectives of the previous experiments.

The data for familiar, distractor, and novel words were analyzed using repeated measures ANOVA, focusing on two factors: object type (same or different) and cue type (pointing or no pointing). No participants were excluded from the analysis, although the average reaction times for familiar words were significantly higher compared to Experiment 4 and previous experiments (See Fig. 27 and Fig. 31). Despite this, the reaction times for all 49 participants were under 1.8 seconds (See Fig.31), leading to the inclusion of all participants in the data analysis.

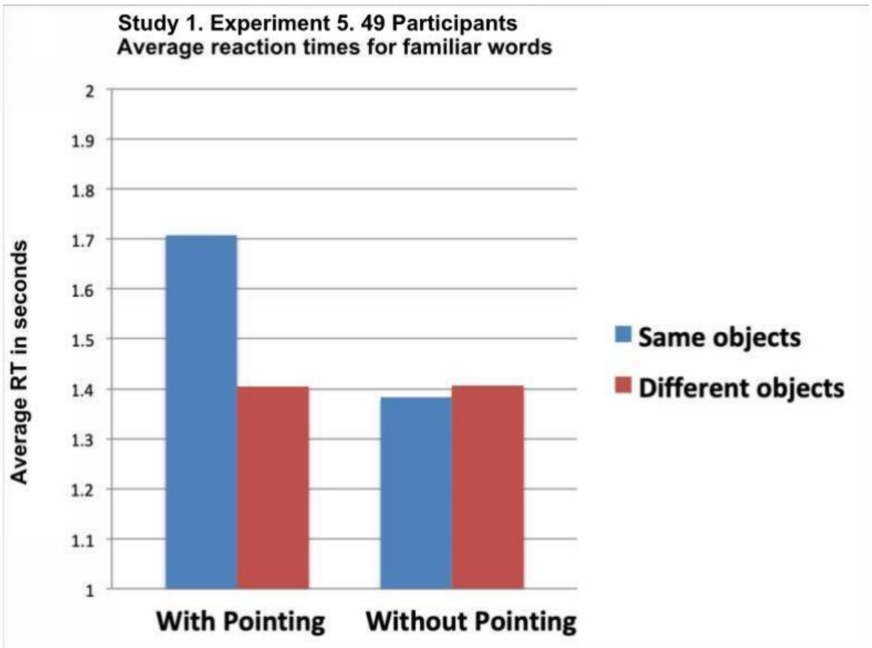



*Figure 31. Experiment 5 : Average reaction times in seconds for familiar words*

# *Results*

The repeated measures ANOVA for the familiar words group showed no significant interaction between Object and Cue ($F(1,48) = .250$, $p = .619$), no main effect of the Object type ($F(1,48) = 1.135$, $p = .292$), and no main effect of the Cue ($F(1,48) = .078$, $p = .781$). However, for the distractor words group, there was a marginal main effect of the Object type ($F(1,48) = 3.327$, $p = .074$) and a marginally significant interaction between Object type and Cue ($F(1,48) = 3.702$, $p = .060$). No main effect of the Cue was found ($F(1,48) = .246$, $p = .622$). Given the marginally significant effects, we proceeded with a Bonferroni-corrected paired samples t-test. For the familiar words group, no significant differences were found between any of the conditions (See Fig.32.). However, for the distractor words group, significant differences were found between the Same Objects/Pointing and Different Objects/Pointing conditions ($p = .015$, 95% C.I. = -.01998, .17390) and between the Same Objects/Pointing and Same Objects/No Pointing conditions ($p = .064$, 95% C.I. = -.00384, .13139) (See Fig.33.). There were no significant differences between the other pairs (See Fig.32 and Fig. 33.).

Finally, a one-sample t-test showed that novel word rejections significantly differed from chance level ($p < .000$) (See Fig.34.).

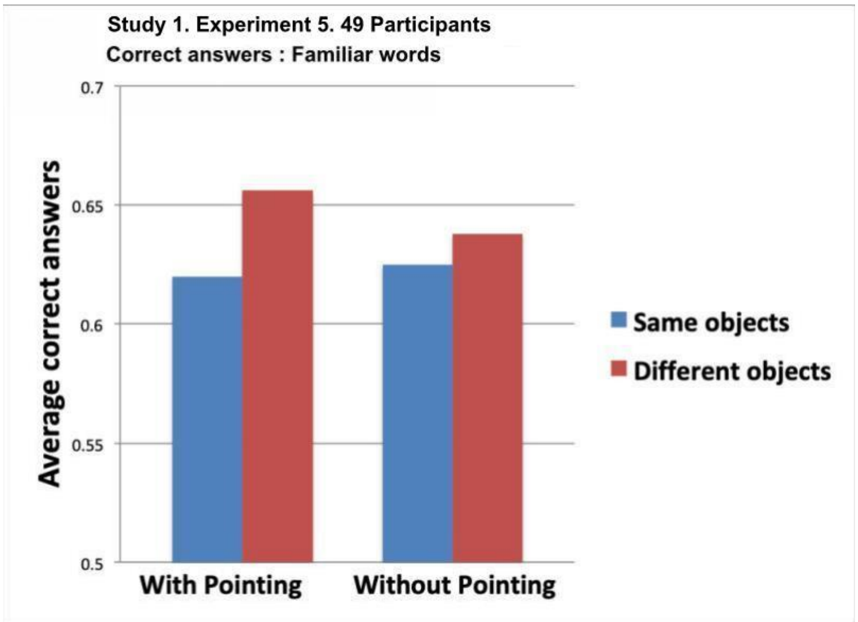



*Figure 32. Experiment 5: Proportion of correct answers for the familiar words in the four conditions*

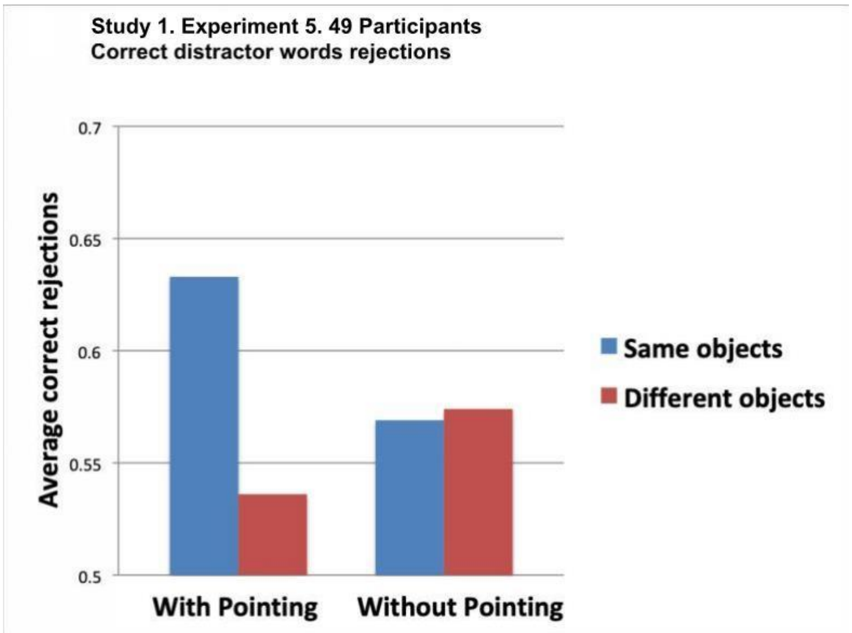



*Figure 33. Experiment 5 : Proportion of correct rejections of the distractor words in the four conditions*

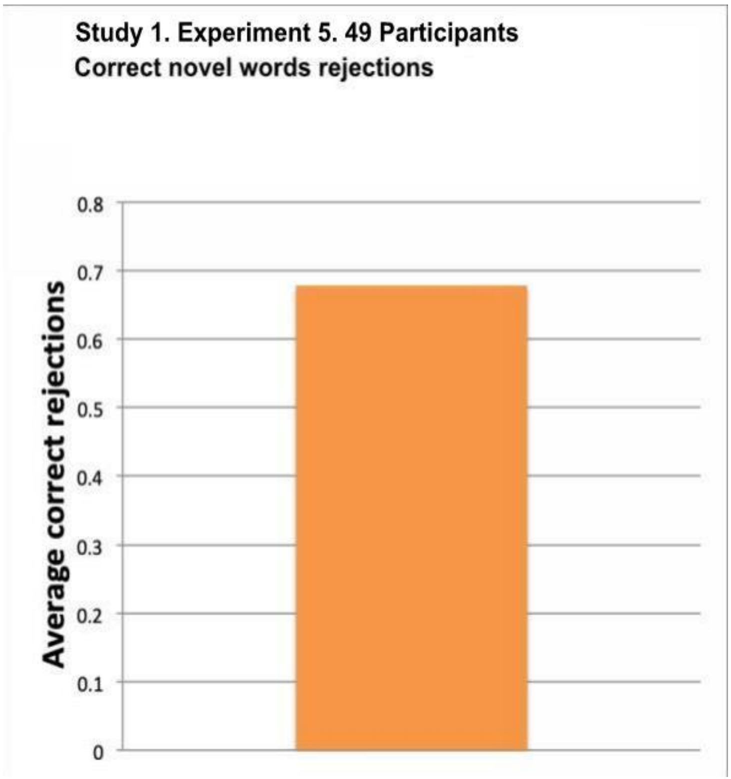



*Figure 34. Experiment 5: Proportion of correct rejections of the novel words*

# *Discussion*

The goal of Experiment 5 was to assess the differences in encoding and memory based on the correctness of answers and varying levels of referential certainty across four conditions, mirroring the objectives of the previous experiments.

In summary, Experiment 5 did not support our hypothesis for the familiar words group but revealed significant effects in the distractor words group. These findings suggest that while pointing gestures and object identity might have slightly influenced the rejection of distractor words, they did not significantly affect the encoding of familiar words in Experiment 5 at least with our sample and in an online context.

## 2.6 *Overall Discussion of Study 1*

Study 1 encompassed five distinct experiments, each designed to explore how varying the levels of referential certainty impact memory encoding in novel word learning scenarios. The experiments employed different learning sessions, along with different cueing mechanisms, to test the hypothesis that higher referential certainty facilitates better encoding. The test phase of each experiment was a simple re-call test of familiar, distractor and novel words.

Experiment 1 set the foundation for the study, hypothesizing that encoding would be more effective under conditions of higher referential certainty. While the analysis revealed no significant effects for object type or cue, a trend toward an interaction between these factors suggested a potential, though subtle, influence of referential certainty on encoding. This subtle effect was supported by a significant difference in the correctness of answers between Same Objects/Pointing and Different Objects/Pointing conditions for familiar words. These result supports our hypothesis, showing that higher levels of referential certainty, enhanced by different objects and pointing gesture to one of them, facilitates better encoding. Still, the results have not been consistent throughout all conditions within Experiment 1. Furthemore, memory was significately better in a conditions with referential cues, no matter the object type. These findings only highlight the importance of social context and social cues in language learning, supporting social- pragmatic theories.

Conducted online, Experiment 2 sought to replicate and extend the findings of Experiment 1, with the reduced complexity of the task and in an online testing enviroment, due to Lockdown because of the Covid 19 pandemic. The lack of significant findings challenged the initial hypothesis. However, reaction times and the consistency in novel word rejection across conditions confirmed that participants were still engaged with the task, although working from environment of their choice.

Experiment 3 aimed to uniformly add an additional referential cue (initial inspection of objects) to all experimental conditions. The results showed again that memory was better when the initial inspection was followed by a pointing gesture, which was true for both familiar words and the rejection of distractor words, no matter the object type. Experiment 4 further investigated the nuances of referential certainty with improved stimuli, according to implications of previous experiments (reduced number of learning sessions and simpler recall-test phase, initial inspection of objects throughout all conditions, corrected object size and duration of learning sessions).

Experiment 5 replicated Experiment 4 in an online environment. Although the results were largely non-significant, the trends and marginal effects suggested that referential certainty played a role, particularly in conditions where pointing cues were used with different objects.

In conclusion, Study 1 has provided interesting but mixed insights into the complex dynamics of referential certainty and its impact on novel word learning. Across all experiments, the performance and memory were generally better in pointing conditions, underscoring the importance of social cues in language learning. Our experiments showed that referential certainty can influence memory encoding particularly when cues are clear and explicit.This led us to rethink our paradigm and make some alterations for the following Study 2.

First of all, we have revaluated the general instruction for the participants in the first five experiments. Participants in Study 1 were instructed to pay attention to the objects in the videos, which may have taken participants' attention from the labeling events. If our instruction was insufficient to trigger associations between the labels and the objects presented, poor learning performance may have resulted. Therefore, we hypothesized that explicit instructions to pay attention to the labels rather than on objects could improve memory performance and enhance the impact of referential certainty on learning. To provide more explicit and straightforward instructions, we considered asking participants to actively engage in learning a new "alien" language, with familiar objects presented as non-distracting elements within the learning environment.

Artificial languages are sometimes a good alternative to studying the natural learning of real words, especially when modelling children's first language acquisition. Recently, many research studies have investigated different factors that influence the learning of pseudowords and artificial languages. For example, Gupta and Tisdale (2009) investigated the role of phonotactic probability in pseudoword learning and found that category typicality had a significant effect on learning speed and accuracy. Similarly, Ellis and Beaton (1993) investigated the cognitive processes involved in learning vocabulary in an artificial language, highlighting the importance of psycholinguistic determinants. Furthermore, de Diego-Balaguer, Martinez-Alvarez, and Pons (2016) found that temporal attention plays a crucial role in scaffolding language development, facilitating the learning of artificial languages. Likewise, a recent neuroimaging study found that compared to familiar words, pseudowords evoke greater activation in a number of language-related brain areas (Fiebach et al., 2016). Many linguistic features that influence the processing of natural English words (e.g., phonotactic probability, phonological neighborhood density, and morphological complexity) have been found to influence the learning and processing time of spoken pseudowords (Kelley & Tucker, 2002).

What many of these studies often have in common is a clear and straightforward learning situation and instruction to concentrate and focus on learning a „new“ language. Therefore, with our Study 2, we aimed to create a more straightforward learning situation and test whether referential certainty might have an impact on the encoding of novel labels when focusing on the „new language“ (i.e., pseudowords).Our goal was to investigate whether varying levels of referential certainty would influence memory encoding in straightforward pseudoword learning situations with using everyday, familiar objects.

In order to test these assumptions, we decided to carry out a new study, with familiar objects in the naming events and more explicit instruction for the participants. Our main research question, whether referential certainty might have an impact on the encoding of novel labels, stayed the same. Study 2 will be described in subsequent text.

# 3. Study 2

To investigate whether different levels of referential certainty would impact memory encoding in simple pseudoword learning tasks using everyday familiar objects, we modified both the stimuli and the experimental instructions for Study 2. Participants were now instructed to attempt to learn and understand the new 'alien language' presented in the videos. This approach was intended to encourage participants to perceive the pseudowords and their referential relationships to the objects on the screen as components of a genuine language. Additionally, we replaced abstract, unfamiliar objects with well-known, familiar ones. We hypothesized that this change would create a more natural language learning environment, reducing distractions and allowing participants to focus more on the referential relationships between the pseudowords and the objects they represented, varying by condition in terms of certainty. Since within Study 1 we introduced two different  testing environments (online and in-person) and they often made a difference, we have decided that Study 2 will also contain two experiments : Experiment 1, tested in-person and Experiment 2, tested online. The further details of Study 2 will be explained in the subsequent text.

## 3.1 Experiment 1

### Stimuli

In creating our new stimuli, we first chose a familiar object base to create our new stimuli and word-learning episodes. We decided to use MultiPic: A standardized set of 750 drawings with multilingual norms offered by BCBL Basque Center on Cognition, Brain, and Language (Meyer, A., New, B., Pliatsikas, C., & Smolka, E). This set contained 750 realistic coloured drawings of objects and animals with standardized visual complexity (Meyer et al.) Pictures were formatted and prepared to use in videos. Secondly, we made a selection of 200 familiar objects for our experiment based on the following criteria:

1. Each drawing was supposed to represent only one whole object, plant or animal.
2. They were supposed to be very common, every day, and easily recognizable.
3. They were supposed to have a few details.
4. Drawings of objects or animals were not supposed to depict any action or movement.
5. They were not supposed to be similar or to be part of the same, narrow object category. All drawings of objects and animals that could possibly be named with the same label were excluded.
6. Drawings that would often evoke some negative or emotional associations were excluded (for example, drawings of a gun, ambulance, snake, or spider).

After finishing our first selection of 220 familiar objects, we asked five people (three friends of the author of this master thesis and two Central European University laboratory members) to look at the selected objects and label them out loud. Any object or animal which took more than 3 seconds to be named was excluded from the further experiment. This way, we could exclude 28 objects and we got our final selection of 192 familiar objects (See Fig.35. for some examples and Supplementary Material for the pictures of all 192 familiar objects) for our word-learning sessions in Experiment IV. In the further text, they will be called 'Familiar objects.'

*Figure 35. Examples of 'Familiar objects' used in Study 2*

Further, we have once again edited our video material in the VSDC Free Video Editor. The newly selected familiar objects were added to the same word-learning sessions. Word learning scenarios were the same as in Experiments 4 and 5 of Study 1. Each video-word-learning session in both experiments of Study 2 starts with a hand wave and the greeting "Szia"- "Hello," followed by the gaze to the booth objects, which are, at this stage, hidden behind occluders (i.e., initial inspection.). After the occluders move, the person provides a novel word (i.e., pseudoword) from the familiar words group in the presence of two familiar objects (See Fig.36.). Like in all previous experiments, in order to modulate the subjects' certainty about the meaning of the novel words, we vary object identity (by using two identical or two different familiar objects) and use of referential cues (having the speaker pointing at one of the familiar objects or not) (See Fig.36 for some examples). Specific for Study 2 was the changed instruction: participants were instructed to focus and try to learn a new 'alien' language they will hear in the word learning sessions. The object exposure time in both Experiment 1 and 2 within Study 2 was 4s, and each word-learning session lasted 13s.

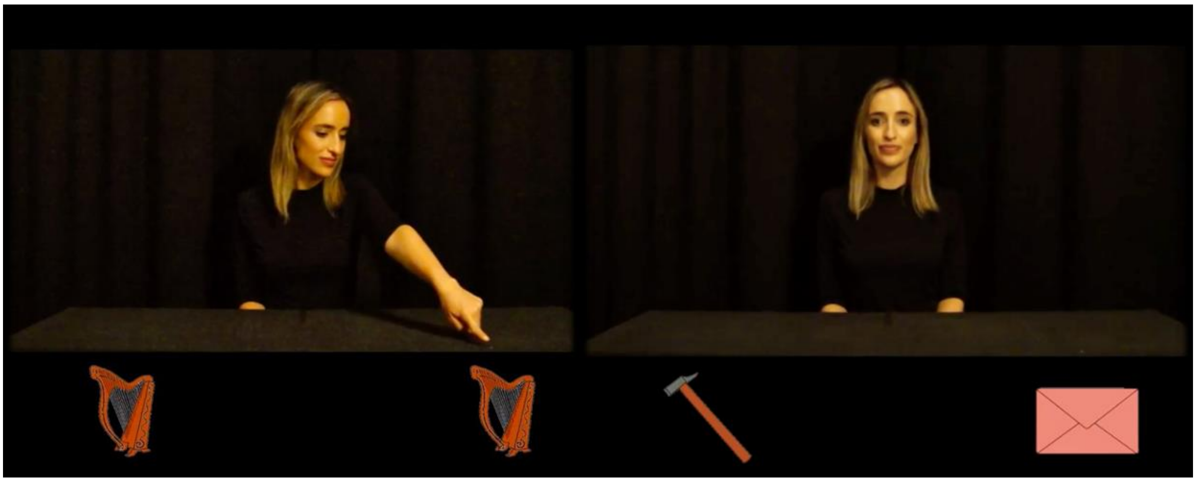


*Figure 36. Experiment IV: Examples of Pointing and No-Pointing conditions with 'Familiar objects'*

# *Experimental Design*

We have designed a new instruction for Study 2 and our new stimuli. Participants were told that they would be presented with word learning videos, where they would hear a new 'alien' language, and that their goal was to try and remember the new words they heard. Specifically, they were instructed to pay attention to the new 'alien language' they will hear in the videos and try to learn and understand this new language as much as possible. In the learning phase of

Experiment 1 and 2 within Study 2, 32-word learning sessions and 32 novel pseudowords from the familiar-word group were presented. Word-learning sessions in four different conditions were played in random order. After watching the movies, the test phase followed. Distractor, novel, and familiar pseudowords were presented on the screen in random order (96 words in total). Like in all previous experiments, participants were instructed to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives.

## *Participants*

In Experiment 1 of Study 2, 40 male and female adults participated, between 18-60 years of age. As in previous experiments, they were all native Hungarian language speakers. Participants were recruited via an advertisement published at Central European University and were rewarded for their time and participation with a small compensation. As in all experiments, the criteria for inclusion were normal hearing, vision, and native knowledge of the Hungarian language.

## *General Procedure*

Participants were sitting in a quiet room at Baby Lab at Central European University. They were greeted, signed the consent form, and started with the experiment. Instructions, the movies of the learning sessions, and the test phase were presented on the computer screen. As in the first and third experiments, learning sessions lasted 13s, and the entire procedure duration was around 30 minutes (depending on the subjects' speed).

Once again, unlike in previous experiments, in Study 2, participants were instructed to focus on the new 'alien language' (rather than on the objects) they will hear in the movies and that their goal is to learn and understand this 'new' language.  Specifically, on the computer screen was written: "You will see short movies about a person speaking in an 'alien' language. Your task is to try to figure out the meaning of the words she is going to say." Our motivation was to direct their attention to the referential relationship between objects and pseudowords they hear.

As our new stimuli, familiar objects in the movies were named with the same pseudowords from the familiar words group in the same word-learning scenario.  Movies of object-naming episodes were presented in random order. During the test phase, once again, participants had to answer the question, "Did you already hear this word?" and they were offered "Yes" and "No" alternatives, and they could indicate their answer by pressing different keys on the keyboard. We collected data about the number of correct answers and subjects' reaction times. The only exclusion criterion was participants' reaction time above 4s, and no participants were excluded in Experiment 1 of Study 2.

## *Data Analysis*

Once again, our experiment aimed to determine whether there are significant differences in the depth of the encoding/memory based on the correctness of the answers and depending on the different levels of referential certainty across four different conditions. In the fourth experiment, the only difference was that our instruction for the participants was more direct (to focus on the 'new' language rather than on objects as in previous experiments). Accordingly, new objects in Study 2 were known and familiar to have a more straightforward learning situation. Still, our hypothesis stayed the same. We expected that the encoding would be better in conditions where two identical familiar objects were shown without a pointing gesture and where two different familiar objects were shown with a pointing gesture toward one of them. In comparison, we expected that the encoding would be worse when two identical familiar objects were shown with a pointing gesture toward one of them and when two different familiar objects were shown without a pointing gesture. The collected data were analyzed using repeated measures ANOVA. We analyzed two factors of object type - with two levels (same or different objects) and cue type with two levels (pointing or no pointing). Our exclusion criteria were participants' reaction time above 4s. No data were excluded in Study 2., Experiment 1 (see Fig. 37. for average reaction time for familiar words in Study 2, Experiment 1).

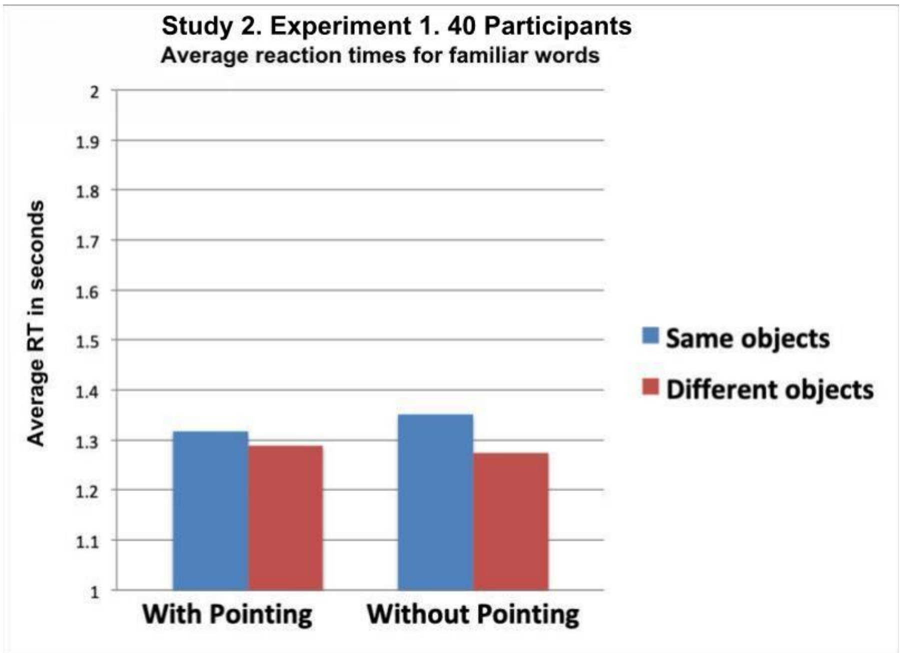



*Figure 37. Experiment 1 , In -person. Average reaction times in seconds for familiar words*

# *Results*

Repeated measures ANOVA was conducted for our In-person version of Experiment IV with 40 participants (N=40). In the familiar words group, we have not found any significant interaction between Object and the Cue F(1,39)=.000 (p =1.000), no main effect of the Object type F(1,39)=.278 (p =.601) or the Cue F(1,39)=.935 (p =.340). Furthermore, in the distractor words group, repeated measures ANOVA have not revealed any main effect of the Object type F(1,39)=.880 (p =.354) or the significant interaction between the Object type and the Cue F(1,39)=.009 (p =.926). We also have not found any main effect of the Cue in the distractor words group F (1,39)=2.604 (p =.115). One sample t-test in the in-person version of Experiment IV showed that novel word rejections significantly differed from the chance level (p<.000) (See Fig.41).

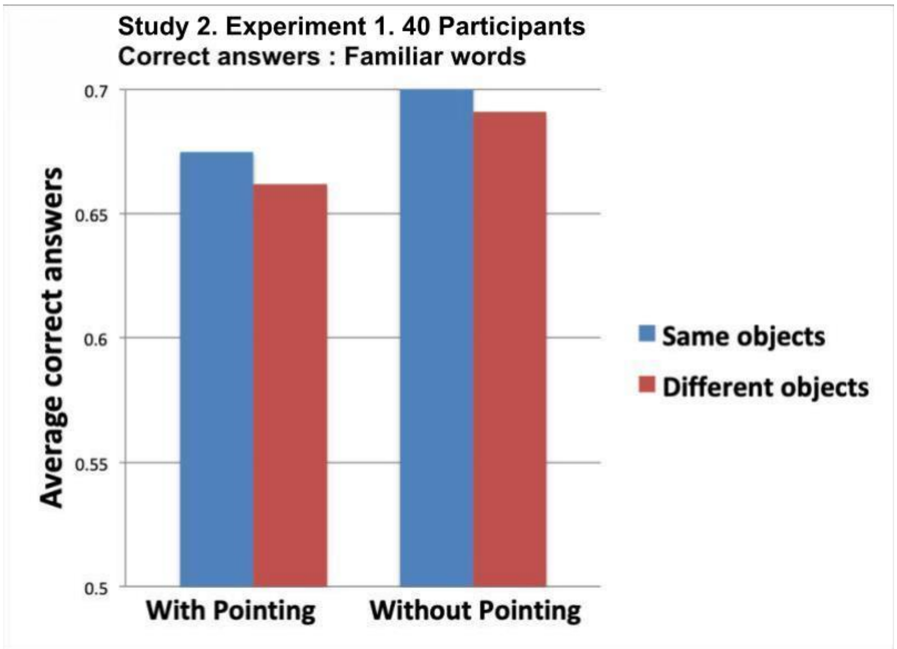



*Figure 39. Experiment 1 : Proportion of correct answers for the familiar words in the four conditions*

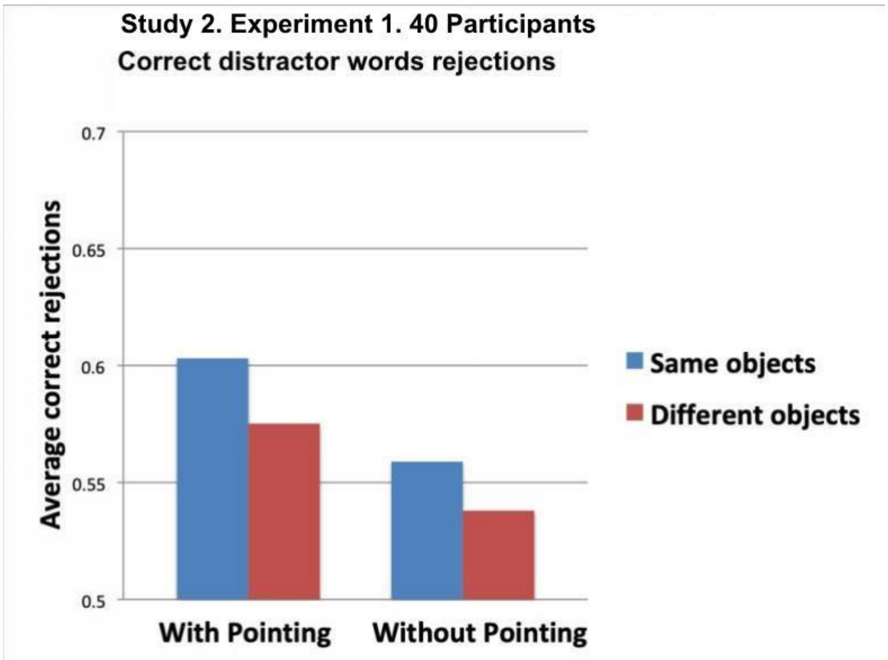



*Figure 40. Experiment 1: Proportion of correct rejections of the distractor words in the four conditions*

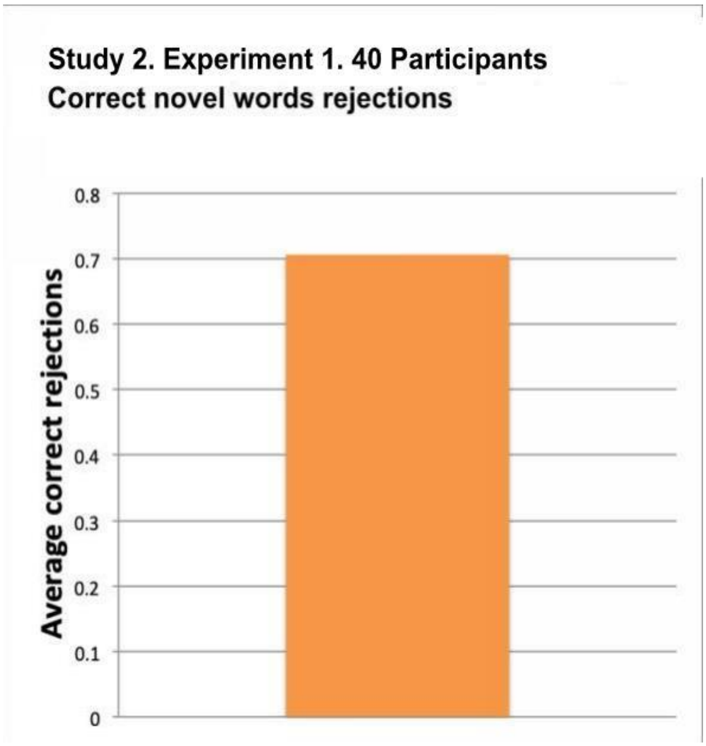



*Figure 41. Experiment 1: Proportion of correct rejections of the novel words*

## *Discussion*

In the Experiment 1 of Study 2, we once again aimed to determine if referential certainty influences memory encoding. Participants were explicitly instructed to focus on learning the new language, a change from previous experiments that focused on objects, and familiar, known objects were presented in learning situations. Our hypothesis remained the same: we expected better encoding in conditions with identical objects and no pointing, as well as different objects with pointing. Conversely, we anticipated worse encoding in conditions with identical objects and pointing, and different objects without pointing.

The results of Experiment 1, however, did not support our hypothesis. The repeated measures ANOVA conducted on the familiar words group with 40 participants revealed no significant interaction between object type and cue, nor any main effect of object type (See Fig.39). Similarly, with the distractor words, there were no significant main effects or interactions: no main effect of object type, no significant interaction between object type and cue (See Fig.40.). Despite these findings, results indicated that novel word rejections significantly differed from the chance level suggesting that participants were indeed paying attention and could distinguish novel words from familiar ones (See Fig.41). In conclusion, we can say that with our sample and experimental conditions, according to results of Experiment 1 of Study 2, we could not confirm our hypothesis. Furthermore, changes made in Study 2, such as more explicit instructions and the use of familiar objects, in the first in person experiment of Study 2 apparently also have not enhanced memory encoding.

## *3.2 Experiment 2*

Since online and in-person testing seemed to make a difference in all the experiments within Study 1, we decided to include online testing in Study 2 as well. The details of online Experiment 2 are described below.

### *Stimuli*

Stimuli was the same as in Experiment 1 of Study 2.

### *Experimental Design*

The experimental design was the same in Experiment 1 and Experiment 2 of Study 2.

### *Participants*

In the Experiment 2 of Study 2, 30 male and female adults participated, again between 18-60 years of age. Experiment 2 of Study 2 was also conducted in early 2021 at Centrale European University in Budapest. All participants were native Hungarian language speakers. They were recruited via the Testable Minds platform (Testable.org) and were rewarded for their time and participation with a small compensation. Once again, the criteria for inclusion were normal hearing, vision, and native knowledge of the Hungarian language.

## *General Procedure*

The Experiment 2 was launched online, on the Testable Minds (Testable.org) platform. Participants were recruited, signed the consent form, and participated in the experiment via the Testable Minds platform. They were working from the environment of their choice, without controlled working conditions. The word learning phase, word recognition test phase, instruction for the participants, and task were the same as in the In-person version of the same experiment.

## *Data Analysis*

As in the first experiment of Study 2, in Experiment 2 of Study 2 we anticipated better encoding when two identical familiar objects were shown without a pointing gesture and when two different familiar objects were shown with a pointing gesture. Conversely, we expected worse encoding when two identical familiar objects were shown with a pointing gesture and when two different familiar objects were shown without a pointing gesture. Data were analyzed using repeated measures ANOVA, considering two factors: object type (same or different) and cue type (pointing or no pointing). Our exclusion criterion was participants' reaction times above 4 seconds, but no data were excluded in Experiment 2 of Study 2 (see Fig. 38 for average reaction times for familiar words in Experiment 2).

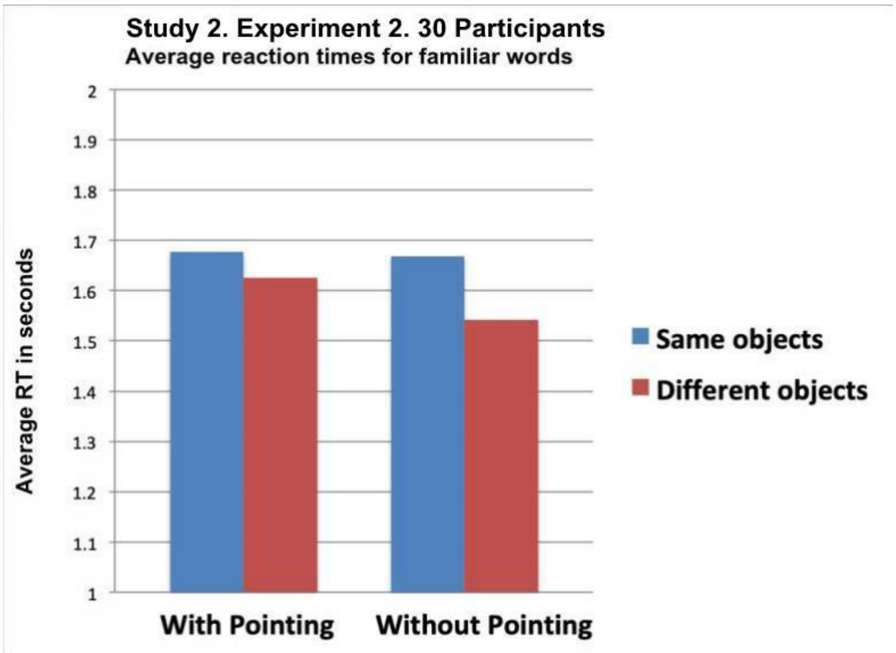



*Figure 38. Experiment 2: Average reaction times in seconds for familiar words*

# *Results*

The repeated measures ANOVA was also conducted for our Experiment 2 (i.e., online version) of Study 2, in which 30 persons participated (N=30). In the familiar words group, the analysis revealed a marginally significant interaction between Object and the Cue F (1,29)=2.962 (p =.096). However, no main effect of the Object type F (1,29)=.442 (p =.511) or the Cue F(1,29)=2.239 (p =.145) were found for the familiar words. For the distractor words, repeated measures ANOVA have not revealed the main effect of the Object type F (1,29)=.301 (p =.588) or the Cue F(1,29)=1.728 (p =.199). Interaction between the Object type and the Cue F(1,29)=.055 (p =.817) was not significant for the distractor words. Because of a marginally significant interaction between Object and the Cue in the familiar words group we have conducted Bonferroni-corrected paired samples t-test, for both familiar and distractor words groups. Bonferroni-corrected paired samples t-test, has revealed a significant difference between the Different Objects/Pointing and Different Objects/No Pointing (*p* = .043, 95% C.I. = .00317, .19683) regarding familiar words (See Fig.42.). The mentioned test has not revealed any significant differences between the Same Objects / No Pointing condition and the Different Objects/No pointing condition (*p* = .419, 95% C.I. = -.05597, .13097), Same Objects/Pointing and Different Objects / Pointing (*p* = .100, 95% C.I. = -.17450, .0616 ), or Same

Objects/Pointing and Same Objects/No Pointing ($p$ = .681, 95% C.I. = -.09865, .06532 ) for the familiar words (See Fig.42). Differences between the same groups regarding the distractor words rejections were not significant (See Fig.43). Furthermore, one sample t-test revealed that distractor words rejections for the Same Objects/No Pointing ($p$ = .126, 95% C.I. = -.0198, .1531 ) and Different Objects / No Pointing ($p$ = .433, 95% C.I. = -.0654, .1488 ) were not different from the chance level, as in the inthe Experiment 1 (See Fig.40 and Fig.43). One sample t-test in the Experiment 2 of Study 2 showed that novel word rejections were significantly different from the chance level (p<.000) (See Fig.44).

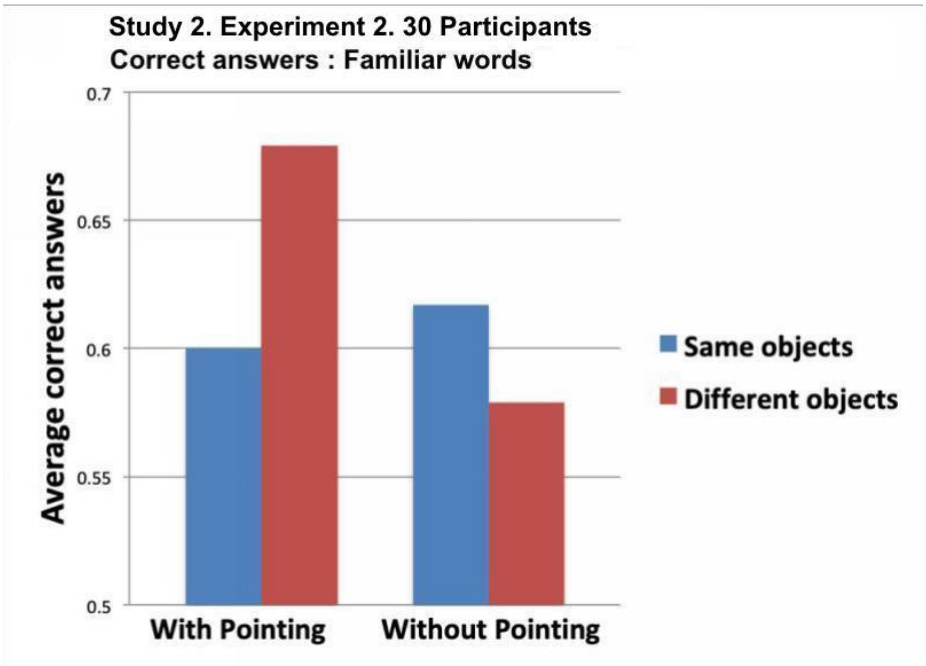



*Figure 42. Experiment 2 : Proportion of correct answers for the familiar words in the four conditions*

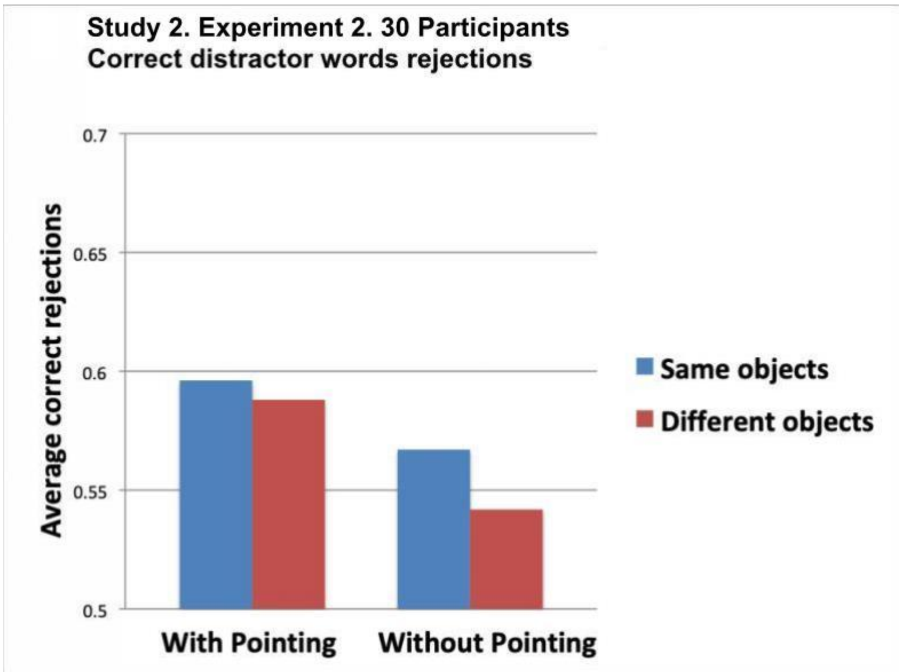



*Figure 43. Experiment 2 : Proportion of correct rejections of the distractor words in the four conditions*

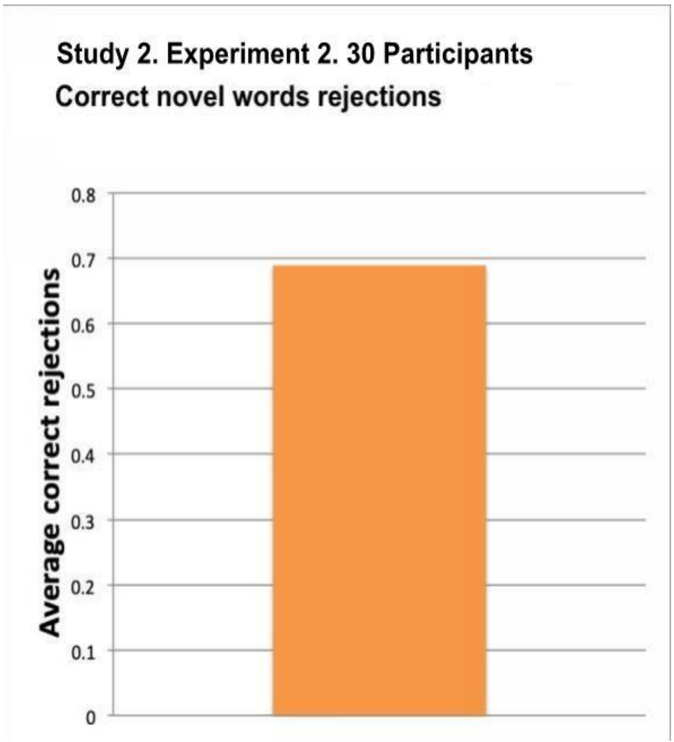



*Figure 44. Experiment 2: Proportion of correct rejections of the novel words*

## *Discussion*

The results of Experiment 2 of Study 2, conducted online with 30 participants, partially confirmed our hypothesis regarding the impact of referential certainty on memory encoding. We anticipated that memory performance would be superior when two identical familiar objects were shown without a pointing gesture and when two different familiar objects were shown with a pointing gesture. Conversely, we expected poorer encoding when two identical familiar objects were shown with a pointing gesture and when two different familiar objects were shown without a pointing gesture.

The repeated measures ANOVA revealed a marginally significant interaction between object type and cue in the familiar words group. This marginal interaction suggests that the combined effect of object type and cue on memory encoding might be more nuanced than initially thought. However, no significant main effects were found for object type. Further analysis using Bonferroni-corrected paired samples t-tests indicated a significant difference between the conditions of different objects with pointing and different objects without pointing in the familiar words group. This finding aligns with our prediction that memory encoding would be better when referential certainty is higher, as evidenced by improved performance in the condition where a pointing gesture clarified the reference of the pseudoword among different objects. In other words, the effect of referential certainty with familiar words was found only when there was a pointing gesture present (See Fig.42). Althlough experiments are quite different, similar effect was found in Experiment 1 of Study 1 (See Fig. 13 ). Once again, these results pinpoin importance of the social cues in making a connection between label and reference.

One-sample t-tests showed that distractor word rejections in the Same Objects/No Pointing and Different Objects/No Pointing conditions were not different from the chance level (See Fig.43). This outcome also mirrors findings from some experiments, stating that the expected improvements in memory encoding due to referential certainty were not uniformly observed across all conditions.

In addition, one-sample t-tests confirmed that novel word rejections were significantly different from the chance level (See Fig. 44), indicating that participants were able to distinguish novel words effectively, despite the nuances in referential certainty conditions.

In summary, Experiment 2 of Study 2 demonstrated that referential certainty, particularly through the use of pointing gestures with different objects, can enhance memory encoding for familiar words. However, this effect was not consistently observed across all conditions, and the distractor words did not show significant differences. These findings suggest that while referential certainty plays a role in memory encoding, according to results of Experiment 2 Study 2, we are still not able to confirm our hypothesis.

## *Overall Discussion of Study 2*

Study 2 aimed to investigate the impact of referential certainty on memory encoding in a pseudoword learning task, building on previous research that explored factors influencing the learning of pseudowords and artificial languages. Prior studies, such as those by Gupta and Tisdale (2009), Ellis and Beaton (1993), and de Diego-Balaguer et al. (2016), have highlighted the importance of various linguistic and cognitive factors, such as phonotactic probability, category typicality, and temporal attention, in the process of learning artificial languages. Inspired by these findings and implications of our Study 1, we sought to create a more ecologically valid and straightforward learning environment. We did that by using familiar, everyday objects and clear instructions to learn a new "alien" language, thus reducing the complexity and ambiguity associated with abstract objects.

In the first experiment of Study 2, participants were tested in-person under controlled conditions. The experimental design incorporated familiar objects instead of abstract ones, and the participants were explicitly instructed to focus on learning a new "alien" language. We hypothesized that memory encoding would be enhanced in conditions where referential certainty was higher—specifically, when two identical familiar objects were shown without a pointing gesture and when two different familiar objects were shown with a pointing gesture toward one of them. Conversely, we expected poorer encoding when referential certainty was lower. However, the results of Experiment 1 did not support our hypothesis. Repeated measures ANOVA revealed no significant main effects of object type or cue type, nor any significant interaction between the four experimental conditions. This lack of significant findings suggests that the modifications made to the experimental setup (using familiar objects and providing more explicit instructions ) may not have enhanced the influence of referential certainty as expected, at least for our sample in an in-person setting.

With the Experiment 2, we replicated Experiment 1 in an online environment. The online version maintained the same experimental design, but participants completed the tasks in a self-selected environment, which as always introduced variability in working conditions and engagement levels. Despite these potential limitations of online testing, the results of Experiment 2 partially supported our hypothesis. Specifically, memory for familiar pseudowords was better in the condition where referential certainty was higher—when a pointing gesture was used to label one of two different familiar objects—compared to when no pointing gesture was used. This finding suggests that referential certainty, facilitated by the combination of familiar objects and pointing gestures, can enhance memory encoding, at least under certain conditions. Moreover, the interaction between object type and cue type was marginally significant, indicating a trend that aligns with our predictions. However, it is important to note that the effects were not consistent across all conditions. In the scenario where two identical familiar objects were present, the use of pointing gestures did not significantly influence memory. In conclusion, the results from Experiment 2 provide partial support for this hypothesis. However, the lack of significant findings in Experiment 1 and the inconsistent effects observed in Experiment 2 suggested once again that that referential certainty, while beneficial, have not uniformly enhanced memory encoding across all learning environments and expected conditions. There is no doubt that relationship between referential certainty and memory encoding is more complex than initially anticipated and that the suggested phenomenon may be sensitive and hard to measure.

Eventually, while Study 2 provides some evidence that referential certainty can enhance memory encoding in pseudoword learning situations with familiar objects, it also underscores the complexity of this relationship and the need for further investigation to fully understand the mechanisms underlying this effect.

# Chapter IV. Final thoughts

## *1. General discussion*

This research project comprised two comprehensive studies, each containing multiple experiments designed to explore the effects of referential certainty on the encoding and memory of novel words in pseudoword learning tasks. Study 1 focused primarily on learning names of abstract objects and varied cueing mechanisms, while Study 2 introduced familiar objects in learning sessions and refined instructional clarity to create a more naturalistic learning environment. Furthermore, we were able to test the desired effect across two working environments: online and in-person.

In Study 1, following Biederman's theory on the primitive components of visual objects, we utilized a set of ten different types of objects, each available in ten different sizes, to build abstract, unfamiliar objects. This approach allowed us to develop 210 novel unfamiliar objects, which we later refined to 192 objects based on our exclusion criteria. Participants in Study 1 learned pseudowords associated with these abstract objects through a series of learning movies. Study 1 encompassed five experiments, all designed to explore how different levels of referential certainty influence memory encoding in pseudoword learning.

Study 1 provided foundational insights, although the results were often subtle and not always statistically significant. In Experiment 1, tested in a labaratory setting, a clear trend toward significance in interaction effects suggested that referential certainty might influence encoding , particularly when distinguishing between different objects in the presence of pointing gesture. In Experiment 2, the first experiment that was tested online, , we did not find any significant differences for the familiar or distractor words, but participants were still able to reliably distinguish novel from familiar words, suggesting effective engagement with the learning task despite the lack of significant findings. Experiment 3 differed from the first two by including a uniform referential cue (initial inspection) across conditions, aiming to to establish a stronger referential connection between the labels and the objects. The results showed a marginally significant main effect of the cue, again particularly in conditions with pointing gesture, supporting the hypothesis that enhanced referential cues could improve memory encoding. Results of Experiment 4 indicated marginal tendencies toward significance in certain conditions, particularly where pointing gestures were involved, highlighting once again the potential contextual dependency of referential certainty effects. Significant differences in distractor word rejections for the pointing conditions suggested that referential cues also play a significant role in the cognitive processing required for distinguishing relevant from irrelevant information.The final experiment of the series, Experiment 5, conducted online, found significant effects primarily for the distractor words, and again better performance in pointing conditions.

The marginal trends and significant differences we found suggested that while referential certainty could enhance memory encoding, their effectiveness was highly context-dependent and is mostly present only in the context with clear referential cues (i.e., pointing gesture). Across all experiments, performance and learning were significantly better in conditions where both referential certainty was higher and pointing gesture was present. When there was no pointing, the levels of referential certainty did not matter as much. Our findings unquestionably highlight the importance of the social cues in the learning situations.

Furthermore, results also revealed the challenges of creating a learning environment for eliciting such subtle effects. The findings of Study 1 identify the need for more straightforward instruction and a more ecologically valid learning environment. Study 2 responded to the ambiguities of Study 1 by employing familiar objects and more explicit instructions aimed at improving the ecological validity of the learning scenarios. Instruction to focus and remember pseudowords and use of everyday objects aimed to create a more straightforward learning environment. The study consisted of two experiments, tested online and in-person. The results from Study 2 were more promising, indicating that the use of familiar objects and clear instructions could indeed enhance the learning experience, as evidenced by improved performance in specific conditions, particularly in online setting (Experiment 2 of Study 2). The analysis of Experiment 2 within Study 2 showed a marginally significant interaction between object type and cue in the familiar words group, suggesting a nuanced effect of these variables on memory encoding. Notably, further analysis revealed a significant difference in memory performance between the conditions of different objects with and without pointing, aligning with the hypothesis that higher referential certainty—achieved through pointing and two different objects present—enhances memory encoding. Again, the observed influence of referential certainty was the most pronounced in conditions with two different objects and pointing gestures toward one of them, both with familiar words and distractor word rejections. We have not found a significant effect in conditions where there was no pointing gesture, no matter the level of referential certainty or object identity. Interestingly, although in different experimental settings and with different stimuli and complexity of the task, results were similar to those in Experiment 1 of Study 1. The importance of clear social cues in word learning was demonstrated once again. In Experiment 1 of Study 2, tested in a laboratory setting, we did not find significant differences between conditions, but performance was still better in pointing conditions and novel word rejections were above chance level, indicating that participants were engaged with the task and could effectively differentiate novel from familiar words. In general, throughout all experiments, enhanced referential cues generally led to improved encoding performance, especially when cues clearly delineated object relationships.

All our experiments in both studies demonstrated that novel word rejections were significantly different from the chance level, indicating consistent engagement with the task across different conditions and settings. Still, across both studies, we could observe the contrast between in-person and online experimental conditions. The main difference was that all online experiments showed higher participants` reaction times and accordingly longer durations of experiment in general. Nonetheless, in all online experiments, participants' success in rejecting distractor and novel words consistently remained above chance level and was not significantly lower than in the laboratory setting. This finding suggests that, despite the uncontrolled online environment, participants maintained high levels of engagement and concentration. Additionally, the generally higher reaction times observed in the online experiments may even indicate greater cognitive effort devoted to recalling each word accurately.Online settings, despite their inherent variability, showed potential and many positive sides, suggesting that digital learning environments can be effectively utilized if designed thoughtfully. Still, the controlled environment in a laboratory settingdid support better performance in general. However, the online environment, despite its less controlled nature, in some cases seemed to allow for a clearer demonstration of the effects of referential certainty (for example, as in Experiment 2 of

Study 2). Comparison of online and laboratory settings across our experiments offered intriguing insights into how environmental factors can influence cognitive tasks.

## *2. Limitations and future directions*

We have three general concerns in the present project. The first important potential research limitation could arguably be found in the empirical study. According to the tendency our results showed, referential certainty is a possible factor influencing word learning. However, such an effect is relatively small, susceptible to alterations, and hard to measure. Furthermore, since no studies have investigated this exact question, in creating our paradigm, we could not find comparable examples. Therefore, our chosen paradigm and developed method may not be sensitive enough and suitable to investigate the desired phenomenon.

Secondly, the results of our experiments suggest that our task was possibly too difficult and that participants' concentration and attention decreased during the repetitionary task. Even after making our task easier after Experiment 1 ( by reducing the number of learning sessions and pseudowords in the test phase ), we have noticed the tendency for longer average reaction times approaching the end of the testing phase, possibly due to repetitiveness and the lack of participants' concentration and engagement. A similar effect was observed in online sessions, throughout the whole experiment, probably due to a less controlled working environment. Therefore, in proceeding with similar experiments, we would suggest a simpler and more actively engaging experiment. For example, investigating object-label associations with fewer repetitions could be a good way to look for further answers.

Finally, similar to most empirical research studies conducted in the laboratory setting, the third significant limitation of our research project is its ecological validity, despite the changes we tried to achieve in Study 2. Although evidence obtained from naturalistic observations of spontaneous speech can often be hard to control and ambiguous (Morgan & Newport, 1981), the language embedded within coherent and comprehensible social activities strongly supports reference learning (Roy et al., 2015). Exposure to entirely new pseudowords in the presence of both abstract and familiar objects was undoubtedly an uncommon experience for our participants and both laboratory and online settings were not able to fully capture the complexity of real-world language learning environments. Therefore, the question if observed, subtle phenomenon would be engaged in natural language learning and real-life situations (for example, in everyday communication with multiple distractors and foreign language learning) remains open. Future research should consider more ecologically valid settings and explore how these findings translate to naturalistic language acquisition contexts.

Additionally, the mixed results between the in-person and online experiments highlight the importance of considering the testing environment's influence on cognitive tasks. Expanding the research to include diverse environments could provide insights into the universality of the findings and highlight variations in cue processing and memory encoding. Furthermore, future research should explore the conditions under which referential certainty is found to be the most effective, possibly by varying the degree of referential ambiguity or by investigating other factors that may interact with referential certainty, such as pseudowords-object associations. Finally, we propose that investigating the suggested phenomenon with first-language learners may be a very inspiring future direction.

# 3. Conclusion

In our empirical study, we intended to elicit different levels of referential certainty. We assumed that using or not using referential cues and gestures in our intentional acts during word learning sessions would make some objects more easily (i.e., with more certainty) selected and remembered as referents for our referential terms. Therefore, we leaned on positive views of reference described by Michaelson and colleagues, especially on the intentionalist model. The intentionalist model describes how references are uttered as part of complex intentional acts. Part of these intentional acts are our gestures, which make specific objects salient in a context (Michaelson et al., 2022). The basic idea is that the speaker's intentions to get the listener to identify a particular object as the referent could be a key to the connection between a term and its reference. These intentions aim to get the listener to identify some object as the referent and suggest how she intends to act (Michaelson et al., 2022). Our results supported this theoretical view. Across all experiments, encoding was more effective in conditions with a pointing gesture. For example, in a condition where different objects were presented with a pointing gesture toward one of them, higher referential certainty was expected and influence on memory was found. In general, across both studies and experiments, our results clearly showed the tendency for more efficient learning when having straightforward referential gestures (i.e., pointing) in the learning sessions.

These findings also support the sensitivity to social cues theory, which asserts that humans vigorously use social-referential information to aid in word learning (Bloom & Markson, 1998; Golinkoff et al., 2000; Tomasello, 2000). In contrast, throughout our study, even if higher referential certainty was expected, we often have not found significant differences in conditions where there was no pointing (even in conditions where an initial look at objects, as an additional referential cue, was present). For example, in conditions where high levels of referential certainty were hypothesized but there was no pointing, (i.e., condition with identical objects and no pointing gesture), the results were much more variable and learning was less successful.

Furthermore, our research underscores the importance of clear and consistent referential cues in teaching new vocabulary. Incorporating gestures and minimizing ambiguity can enhance learners' encoding and retention of new words, making the learning process more effective. The effectiveness of learning new pseudowords in our Study 2, especially Experiment 2, specifies the importance of clear and straightforward instruction in learning situations. Again, the learning was the most successful in the condition where two different objects were present and there was a pointing gesture toward one of them; in other words, in a condition where referential certainly was high and there was a pointing gesture. There is no doubt that the complex interplay between referential certainty and other cognitive factors during word learning is supported by clear and intentional referential cues and straightforward learning instructions, so it would be advisable to consider these factors in future investigations and learning environments.

Throughout both studies and seven experiments, we also had an opportunity to experience and compare two testing environments: in-person and online. Even if the lockdown in Europe during Covid 19 Pandemy was our first motivation to create online experiments, we have recognized many benefits of online testing in practice (less time and space, easier participant recruitment, bigger samples, and quicker and more efficient testing). Online settings, despite their inherent variability, showed promise especially under conditions of enhanced instructional clarity and object familiarity (Study 2, Experiment 2), suggesting that digital learning environments can be effectively utilized if designed thoughtfully.

This project has highlighted the nuanced ways in which referential certainty, object familiarity, and instructional clarity can impact the learning of new languages. For educators, the results underscore the importance of using relatable materials and clear, direct instructions to enhance learning. For experimental psychologists, the findings call for more nuanced designs that consider the interplay of environmental variables and instructional clarity.

In conclusion, while our results provide only partial evidence that referential certainty directly influences the encoding of novel words, they offer a valuable starting point for future research. Our findings suggest that while referential certainty plays a role in memory encoding, other factors also contribute to the complexity of this process. Further research is necessary to delineate these contributing factors and refine our understanding of how referential certainty influences language learning. Although according to our results we cannot claim that our paradigm is sensitive enough to investigate this phenomenon, we believe that the developed material made solid ground for anyone interested in investigating this relatively unexplored question. Our paradigm has the potential to be assembled further and to shed light on the general relationship between memory encoding and certainty about the meaning of incoming information. Future studies should build on our findings to explore this relationship further, considering additional variables, as object-referent associations, and longer-term effects. Furthermore, we propose that investigating influence of referential certainty with first-language learners (i.e., children) may be a very inspiring future direction. By continuing to explore these dynamics, future work can help optimize learning environments and instructional methodologies, potentially transforming practices in educational settings and beyond.

## *Supplementary Material*

## *Instructions and consent form*

## 1. Consent form (original):

## Tájékoztató nyilatkozat
## Kedves Résztvevő!

Köszönjük, hogy részt vesz a kutatásunkban! Ebben a kísérletben azt vizsgáljuk, hogy milyen tényezők befolyásolják az emlékezeti kódolást új ingerek esetén. A vizsgálat során rövid filmeket fogunk mutatni, amikben egy személy minden alkalommal két ismeretlen tárgyat mutat, illetve mond egy ismeretlen szót. Ezek a tárgyak bizonyos esetekben azonosak lesznek, bizonyos esetekben pedig különbözni fognak egymástól. Az Ön feladata az lesz, hogy próbáljon meg minél jobban odafigyelni és megjegyezni ezeket a tárgyakat. Miután végignézte ezeket a filmeket, egy emlékezeti tesztet kell kitöltenie, ahol a válaszait két billentyű lenyomásával tudja majd megadni.

Minden általunk rögzített adatot bizalmasan kezelünk. Csak azoknak a személyeknek lesz hozzáférésük a felvételekhez, akik a kutatásban dolgoznak. A vizsgálatban a résztvevők egy-egy kódszámot kapnak, így az adatokból az egyén nem azonosítható. A vizsgálat alapján készült tudományos tanulmányok és kutatási beszámolók csak a csoport átlagteljesítményét fogják tartalmazni, az egyes résztvevők adatait nem. A vizsgálat végeztével amennyiben erre igényt tart, beszámolunk arról, hogy mit tudtunk meg a kutatásból a vizsgált csoport eredményei alapján.

Ha bármilyen kérdése van a vizsgálattal vagy a módszereinkkel kapcsolatban, megkeresheti a vizsgálat vezetőjét, Dr. Marno Hannát a hanna.marno@gmail.com email címen, illetve a +36 30 2449832 telefonszámon. A vizsgálatot Danyi Róbert végzi, akihez közvetlenül a danyir@ceu.edu email címen fordulhat.

Kérjük, amennyiben beleegyezik, aláírásával igazolja, hogy elolvasta az ismertetőnket, és hozzájárul a vizsgálatban való részvételéhez. Hozzájárulását, amelyet ezen a papíron megerősít, bármikor, következmények nélkül visszavonhatja.

Neve:

Születési dátuma:

Aláírása:

Dátum:

## 2. Consent Form (translation):

Dear Participant,

Thank you for taking part in our study! The goal of our research is to study those factors that might have an impact on encoding novel stimuli. During the study, we are going to present short movies in which each time a person is going to show two unfamiliar ob-jects and say an unfamiliar word. Sometimes these objects will be identical, and some-times they will be different. Your task will be to pay attention as much as you can and remember these objects. After seeing the movies, you will be asked to fill a memory questionnaire by pressing two keys on your keyboard.

All the data we record will be treated confidentially. Only those people can access the data who work in the project. All participants will receive a code, in this way participants cannot be identified. All the scientific reports and presentations based on this study will include the average performance of the group, and not the individual data of the partic-ipants. Once the study is finished, upon request, we can inform you about the conclu-sions of the group results.

In case you have any question about the study, don't hesitate to contact the leader of the project, Dr Hanna Marno (email: hanna.marno@gmail.com, phone: 0036302449832). The study is conducted by Robert Danyi, who can be contacted via email (danyir@ceu.edu).

In case you agree to participate in the study, please provide your signature to confirm that you have read the information and that you are willing to participate in the study. Any point during the study you can withdraw your participation without consequences.

Name:

Date of birth:

Signature:

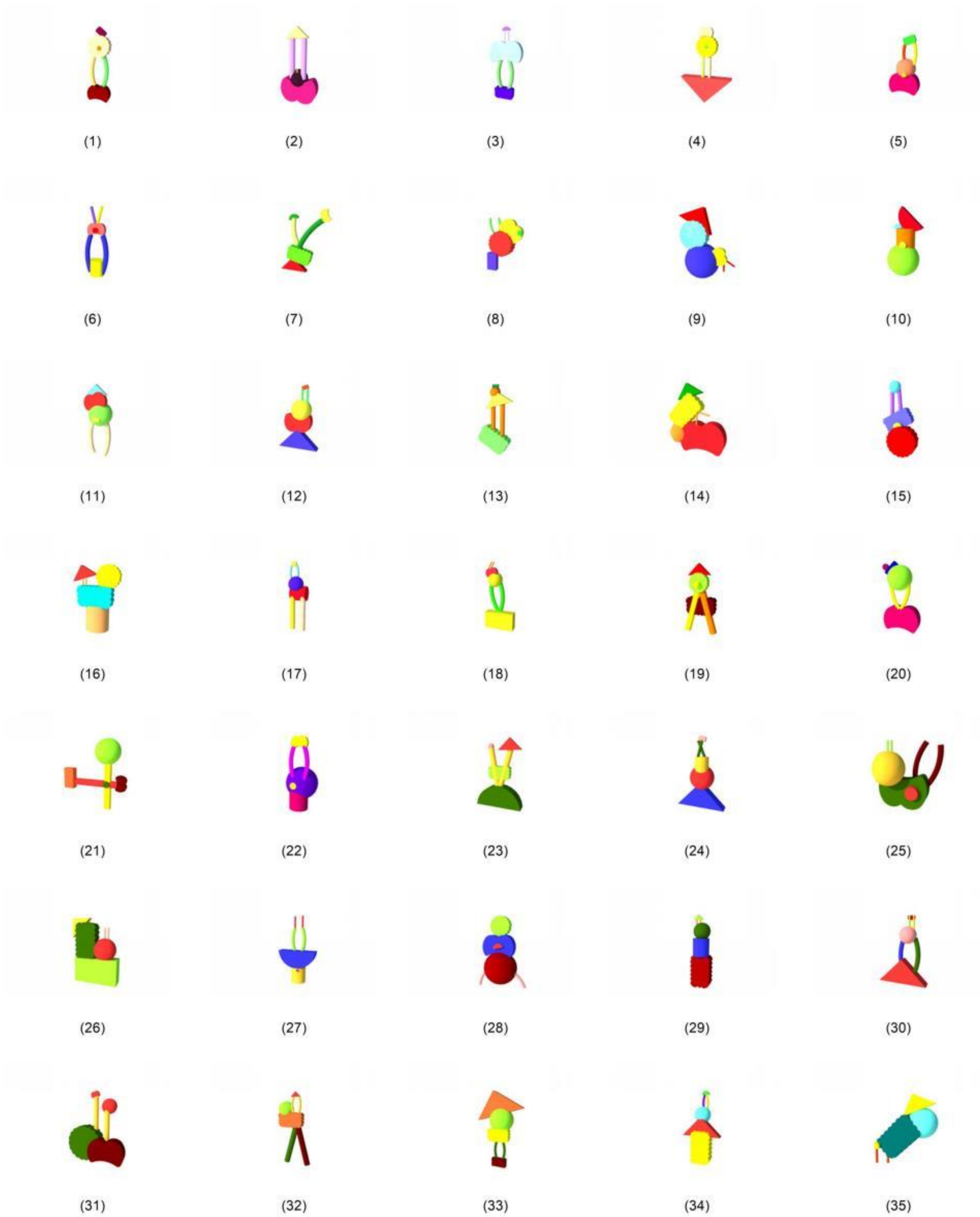
Date:

## 3. Novel (Pseudo) Words and their Distractor- Word Pairs

| Novel Words | Distractor Words |
|---|---|
| NACSU | DACSU |
| MODI | LODI |
| BESU | BETU |
| CIBA | CILA |
| DÁGU | BÁGU |
| MIRU | MIBU |
| JÜVÉ | JÜBÉ |
| FIMU | FILU |
| GOPA | NOPA |
| HELU | BELU |
| DIME | SIME |
| JUMA | JUBA |
| KÜSE | KÜTE |
| LÁSU | PÁSU |
| MURÓ | VURÓ |
| NESU | NELU |
| TIVÓ | KIVÓ |
| PUGÓ | PUSÓ |
| REVÓ | DEVÓ |
| SÖZÜ | KÖZÜ |
| TAZU | RAZU |
| ZUBÉ | ZUKÉ |
| VECÜ | VEHÜ |
| ZADU | BADU |
| FABÓ | MABÓ |
| PÁTA | FÁTA |
| GIBO | KIBO |
| RICU | RILU |
| DUNI | LUNI |
| ZEGA | TEGA |
| TÉFE | TÉBE |
| SAFI | ZAFI |
| GOBA | GOZA |
| HILU | HITU |
| LEMI | GEMI |
| KEJÓ | ZEJÓ |
| RIKE | RIPE |
| ZÁBA | DÁBA |
| VUZE | KUZE |
| NAKÓ | NABÓ |
| TOGE | POGE |
| SÉBU | TÉBU |
| RÁPU | GÁPU |
| DOFE | ROFE |

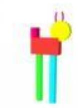
| | |
|---|---|
| POKI | PORI |
| RUDA | DUDA |
| JALÓ | JATÓ |
| BOLI | BOKI |
| DOSU | BOSU |
| GUCI | ZUCI |
| SENU | SEPU |
| BECA | BERA |
| VOGA | MOGA |
| DOSI | KOSI |
| FIHE | FOHE |
| GELA | GILA |
| HUMI | HULI |
| JISU | GISU |
| KUTA | KUGA |
| LODU | PODU |
| MIPÉ | MIDÉ |
| NUPÓ | FUPÓ |
| PEVI | PEZI |
| RATÓ | RABÓ |

# 4. Novel (Abstract) Objects




(36)

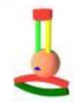
(37)

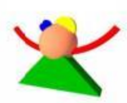
(38)

(39)

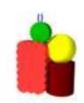
(40)

(41)

(42)

(43)

(44)

(45)

(46)

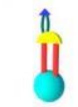
(47)

(48)

(49)

(50)

(51)

(52)

(53)

(54)

(55)

(56)

(57)

(58)

(59)

(60)

(61)

(62)

(63)

(64)

(65)

(66)

(67)

(68)

(69)

(70)


(71)

(72)

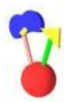
(73)

(74)

(75)

(76)

(77)

(78)

(79)

(80)

(81)

(82)

(83)

(84)

(85)

(86)

(87)

(88)

(89)

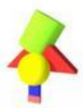
(90)

(91)

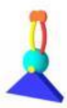
(92)

(93)

(94)

(95)

(96)

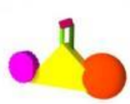
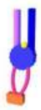
(97)

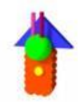
(98)

(99)

(100)

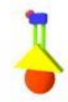
(101)

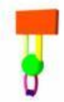
(102)

(103)

(104)

(105)

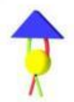

(106)


(107)


(108)


(109)


(110)


(111)


(112)


(113)


(114)


(115)


(116)


(117)


(118)


(119)

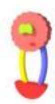

(120)


(121)


(122)


(123)


(124)


(125)


(126)


(127)


(128)


(129)

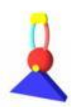

(130)


(131)


(132)


(133)


(134)


(135)


(136)


(137)


(138)


(139)


(140)


(141)


(142)


(143)

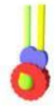

(144)


(145)


(146)


(147)


(148)


(149)


(150)

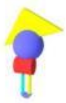

(151)


(152)


(153)


(154)


(155)


(156)


(157)


(158)


(159)


(160)


(161)


(162)


(163)


(164)

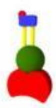

(165)

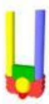

(166)


(167)

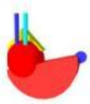

(168)

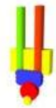

(169)


(170)


(171)

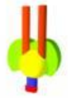

(172)

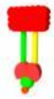

(173)


(174)


(175)


(176)

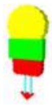

(177)


(178)


(179)


(180)


(181)


(182)


(183)

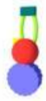

(184)


(185)


(186)


(187)


(188)


(189)


(190)


(191)


(192)

# 5. Familiar Objects


(1)


(2)


(3)


(4)


(5)


(6)


(7)


(8)


(9)

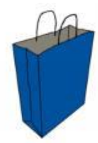

(10)


(11)


(12)


(13)

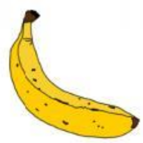

(14)

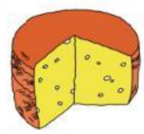

(15)


(16)


(17)


(18)


(19)


(20)


(21)


(22)


(23)


(24)


(25)


(26)


(27)


(28)


(29)


(30)


(31)


(32)


(33)


(34)


(35)


(36)

(37)

(38)

(39)

(40)

(41)

(42)

(43)

(44)

(45)

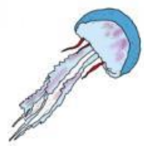
(46)

(47)

(48)

(49)

(50)

(51)

(52)

(53)

(54)

(55)

(56)

(57)

(58)

(59)

(60)

(61)

(62)

(63)

(64)

(65)

(66)

(67)

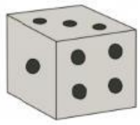
(68)

(69)

(70)


(71)

(72)

(73)

(74)

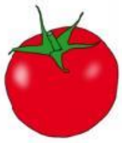
(75)

(76)

(77)

(78)

(79)

(80)

(81)

(82)

(83)

(84)

(85)

(86)

(87)

(88)

(89)

(90)

(91)

(92)

(93)

(94)

(95)

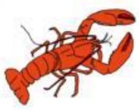
(96)

(97)

(98)

(99)

(100)

(101)

(102)

(103)

(104)

(105)


(106)


(107)


(108)

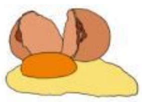

(109)


(110)


(111)


(112)


(113)


(114)


(115)


(116)


(117)


(118)


(119)


(120)


(121)


(122)


(123)


(124)


(125)


(126)


(127)


(128)


(129)


(130)


(131)


(132)


(133)


(134)


(135)


(136)


(137)


(138)


(139)


(140)


(141)


(142)


(143)


(144)


(145)


(146)


(147)


(148)

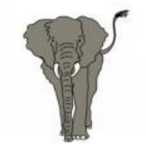

(149)


(150)


(151)


(152)


(153)


(154)


(155)


(156)


(157)


(158)


(159)


(160)


(161)


(162)


(163)


(164)


(165)


(166)


(167)


(168)


(169)


(170)


(171)


(172)


(173)


(174)


(175)


(176)


(177)


(178)


(179)


(180)

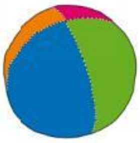

(181)


(182)


(183)

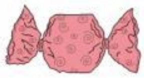

(184)


(185)


(186)


(187)


(188)


(189)


(190)


(191)


(192)

# VI. References


1. Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty in early word learning. Child development, 67(2), 635-645.

2. Allen Preissler, M., & Carey, S. (2004). Do both pictures and words function as symbols for 18-and 24-month-old children?. *Journal of cognition and development*, *5*(2), 185-212.

3. Baldwin, D. A., & Tomasello, M. (1998). Word learning: A window on early pragmatic understanding.

4. Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. Child development, 67(6), 3135-3153.

5. Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. Psychological review, 94(2), 115.

6. Bloom, P., & Markson, L. (1998). Capacities underlying word learning. Trends in Cognitive Sciences, 2(2), 67-73.

7. Brooker, R. J., Buss, K. A., Lemery-Chalfant, K., Aksan, N., Davidson, R. J., & Goldsmith, H. H. (2013). The development of stranger fear in infancy and toddler-hood: Normative development, individual differences, antecedents, and out-comes. *Developmental science*, *16*(6), 864-878.

8. Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict ac-celerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. Journal of child language, 35(1), 207-220.

9. Cartmill, E. A., Armstrong III, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabu-lary 3 years later. Proceedings of the National Academy of Sciences, 110(28), 11278-11283.

10. Chen, D. Y., Tian, X. P., Shen, Y. T., & Ouhyoung, M. (2003, September). On visual similarity based 3D model retrieval. In Computer graphics forum (Vol. 22, No. 3, pp. 223-232). Oxford, UK: Blackwell Publishing, Inc.

11. Chomsky, N. (1957). Logical structure in language. *Journal of the American Society for Information Science*, *8*(4), 284.

12. Chomsky, N. (1959). On certain formal properties of grammars. Information and control, 2(2), 137-167.

13. Chomsky, N., & Halle, M. (1965). Some controversial questions in phonological theory. *Journal of linguistics*, *1*(2), 97-138.

14. Chomsky, N. (2000). New Horizons in the Study of Language and Mind. Cambridge MA: The MIT Press.

15. Chomsky, N. (2012). The Science of Language and Mind. Interviews with James McGilvray. Cambridge MA: Cambridge University Press.

16. Clark, E. V., & Berman, R. A. (1987). Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child language*, *14*(3), 547-567.

17. de Beeck, H. P. O., Torfs, K., & Wagemans, J. (2008). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. Journal of Neuroscience, 28(40), 10111-10123.

18. de Diego-Balaguer, R., Martinez-Alvarez, A., & Pons, F. (2016). Temporal attention as a scaffold for language development. Frontiers in psychology, 7, 44.

19. Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science*, *171*(3968), 303-306.

20. Eisenbeiss, S. (2009). Generative approaches to language learning. Linguistics, 47(2), 273-310.

21. Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. Language Learning, 43(4), 559-617. https://doi.org/10.1111/j.1467-1770.1993.tb00627.x

22. Fiebach, C. J., Friederici, A. D., Müller, K., & Von Cramon, D. Y. (2002). fMRI evidence for dual routes to the mental lexicon in visual word recognition. *Journal of cognitive neuroscience*, *14*(1), 11-23.

23. Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. Child development, 1532-1541.

24. George J.. Hollich, Hirsh-Pasek, K., & Golinkoff, R. M. (2000). Breaking the Language Barrier: An Emergentist Coalition Model for the Origins of Word Learning. University of Chicago Press.

25. 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.

26. González-Fernández, B., & Schmitt, N. (2020). Word knowledge: Exploring the relationships and order of acquisition of vocabulary knowledge components. Applied Linguistics, 41(4), 481-505.

27. Gupta, P., & Tisdale, J. (2009). Word learning, phonological short-term memory, and phonotactic probability: The influence of category typicality on learning speed and accuracy. Journal of Experimental Psychology: Learning, Memory, and Cognition, 35(3), 484-498. https://doi.org/10.1037/a0014878

28. Hockett, C. F., & Hockett, C. D. (1960). The origin of speech. Scientific American, 203(3), 88-97.

29. Howell, D. C. (2012). *Statistical methods for psychology*. Cengage Learning.

30. Jusczyk, P. W., & Hohne, E. A. (1997). Infants' memory for spoken words. *Science*, *277*(5334), 1984-1986.

31. Kaplan, D. (1990). Words. *Proceedings of the Aristotelian society, supplementary volumes*, *64*, 93-119.

32. Kelley, M. C., & Tucker, B. V. (2022). Perception and timing of acoustic distance.

33. Kuhl, P. K. (2004). Early language acquisition: cracking the speech code. Nature reviews neuroscience, 5(11), 831-843.

34. Kuhl, P. K., Williams, K. A., Lacerda, F., Stevens, K. N., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606-608.

35. Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. Science, 330(6012), 1830-1834.

36. Lederberg, A. R., Prezbindowski, A. K., & Spencer, P. E. (2000). Word-learning skills of deaf preschoolers: the development of novel mapping and rapid word-learning strategies. Child development, 71(6), 1571-1585.

37. Markman, E. M. (1987). How children constrain the possible meanings of words.

38. Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. Perspectives on language and thought: Interrelations in development, 72-106.

39. Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. Cognitive psychology, 20(2), 121-157.

40. Marno, H., Davelaar, E. J., & Csibra, G. (2014). Nonverbal communicative signals modulate attention to object properties. Journal of Experimental Psychology: Human Perception and Performance, 40(2), 752.

41. McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. Psychological review, 119(4), 831.

42. Meyer, A., New, B., Pliatsikas, C., & Smolka, E. MultiPic: A standardized set of 750 drawings with multilingual norms.

43. Michaelson, Eliot and Marga Reimer, "Reference", The Stanford Encyclopedia of Philosophy (Summer 2022 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2022/entries/reference/>.

44. Merriman, W. E., Marazita, J., & Jarvis, L. (1995). Children's disposition to map new words onto new referents. In M. Tomasello & W. E. Merriman (Eds.), Beyond names for things: Young children's acquisition of verbs (pp. 149–163). Hillsdale, NJ: Erlbaum.

45. Morgan, J. L., & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of verbal learning and verbal behavior*, *20*(1), 67-85.

46. Plunkett, K., & Sinha, C. (1992). Connectionism and developmental theory. British Journal of Developmental Psychology, 10(3), 209-254.

47. Quine, W.V.O., 1960. Word and Object, Cambridge MA: MIT Press.

48. Reimer, M. (2009). Stanford Encyclopedia of Philosophy:" Reference. Retrieved on April, 4, 2023.

49. Richards, J. C. (1976). The role of vocabulary teaching. TESOL quarterly, 77-89.

50. Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. Proceedings of the National Academy of Sciences, 112(41), 12663-12668.

51. Roy, D. K., & Pentland, A. P. (2002). Learning words from sights and sounds: A computational model. Cognitive science, 26(1), 113-146.

52. Roy, D., Patel, R., DeCamp, P., Kubat, R., Fleischman, M., Roy, B., ... & Levit, M. (2006, September). The human speechome project. In International Workshop on Emergence and Evolution of Linguistic Communication (pp. 192-196). Springer, Berlin, Heidelberg.

53. Senju, A., & Csibra, G. (2008). Gaze following in human infants depends on communicative signals. Current Biology, 18(9), 668-671.

54. Sloutsky, V. M., Lo, Y. F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child development*, *72*(6), 1695-1709.

55. 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.

56. Sloutsky, V. M., & Robinson, C. W. (2008). The role of words and sounds in infants' visual processing: From overshadowing to attentional tuning. *Cognitive Science*, *32*(2), 342-365.

57. Smith, L. B. (2000). How to learn words: An associative crane. Breaking the word learning barrier.

58. Swingley, D. (2009). Contributions of infant word learning to language development. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1536), 3617-3632.

59. Todorović, D. (1995). *Osnovi metodologije psiholoških istraživanja*. Laboratorija za eksperimentalnu psihologiju.

60. Trask, R. L. (2014). Introducing linguistics: a graphic guide. Icon Books Ltd.

61. Von Humboldt, W. (1999). *Humboldt:'On language': On the diversity of human language construction and its influence on the mental development of the human species*. Cambridge University Press.

62. Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in cognitive sciences*, *13*(6), 258-263.

63. Waxman, S. R., & Kosowski, T. D. (1990). Nouns mark category relations: Toddlers' and preschoolers' word-learning biases. Child development, 61(5), 1461-1473.

64. Woodward, A. L., & Markman, E. M. (1998). Early word learning. In D. Kuhn & R. S. Siegler (Eds.), W. Damon (Series Ed.), Handbook of child psychology: Vol 2. Cognition, perception, and language (5th ed., pp. 371–420). New York: Wiley

65. Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: Salience and social cues in early word learning. *Developmental Science*, *20*(2), e12349.