UNIVERSITÄT DES SAARLANDES
PHILOSOPHISCHE FAKULTÄT II
SPRACH-, LITERATUR- UND KULTURWISSENSCHAFTEN

The Interactive Nature of Second-Language Word Learning in Non-Instructed Environments

Dissertation

zur Erlangung des akademischen Grades eines

Doktors der Philosophie

der Philosophischen Fakultäten

der Universität des Saarlandes

vorgelegt von

Judith Köhne

aus Recklinghausen

Saarbrücken, 2011

Dekan:	Prof. Dr. Erich Steiner
Berichterstatter/innen:	Prof. Dr. Matthew W. Crocker
	Prof. Dr. Martin J. Pickering
Tag der letzten Prüfungsleistung:	22.11.2011

Abstract

Gaining the command of a second language is a difficult task for an adult. Understanding and learning novel words is challenging, particularly in non-instructed situations: Words are often parts of complex linguistic contexts and potential referents are embedded in rich visual scenes. To overcome this challenge learners can potentially exploit the richness of their multi-modal environment through a range of different word-learning mechanisms and based on automatic sentence-processing mechanisms.

Despite numerous investigations of word learning by researchers from a range of disciplines, few have examined the interplay of different learning and processing mechanisms. Such an approach, however, potentially both oversimplifies and overcomplicates the scenario. Moreover, most studies suffer from either a lack of naturalness or a lack of control of the experimental items.

The main goal of this thesis is to study word learning in adults in a more situated and interactive manner, considering different mechanisms, processes, and information sources in parallel. This enterprise is driven by the motivation to contribute to a more complete theory of second-language word learning, to bridge research traditions, and to draw implications for the development of practical learning applications. In particular, we examined the interaction of the two important and visually situated word-learning mechanisms, *cross-situational word learning* (CSWL, Yu & Smith, 2007) and *sentence-level constraint learning* (SLCL). CSWL is a bottom-up, associative manner of word learning: people make connections between visual objects and spoken words by tracking their co-occurrence frequencies. SLCL, on the contrary, is a top-down strategy, which is based on making inferences about likely word meanings given a linguistic context (and word knowledge). SLCL in this thesis refers to inferring the meanings of object nouns (e.g., *the corn*) based on restrictive verbs (e.g., *eat*), a visual context, and people's world knowledge.

Our studies exploit a novel experimental paradigm which integrates teaching German adults a semi-natural miniature language in a step-wise procedure. Participants were familiarized with a set of verbs (e.g., *bermamema* 'to eat') before they were exposed to noun-learning trials. These trials consisted of pairs of visual scenes and auditory transitive sentences, in which novel nouns were embedded (e.g. *Si laki bermamema si sonis.*, 'The man will eat the corn'). Finally, participants performed a forced-choice vocabulary test (with confidence ratings). Eye-movements were recorded during learning and testing.

In Experiment 1, we evaluated the use of CSWL and SLCL in this naturalized situation. We found that participants applied both mechanisms in a complementary manner to learn the vocabulary. Moreover, SLCL (verbal constraints) clearly boosted on-line identification of referents. In Experiment 2, we introduced a second word order (OVS), which is characterized by a verb which follows rather than precedes the syntactic object (that denotes a visual object). Results are in accordance with the hypothesis that verb-based prediction of referents has a positive influence on noun learning. In Experiment 3, we re-addressed the question whether SLCL boosts noun learning and examined the interaction of CSWL and SLCL by manipulating the degree of referential uncertainty. On-line and off-line results provide evidence for the hypotheses that, firstly, SLCL boosts noun learning and secondly, SLCL and CSWL interact in that they jointly contribute to the identification of noun meanings. Experiment 4 was conducted in order to investigate the interaction of CSWL and SLCL when both mechanisms are in conflict. Nouns had two potential meanings with different co-occurrence frequencies. Since the verbal restrictions supported the meaning with the lower frequency, CSWL and SLCL were in conflict. Learning rates clearly reveal that CSWL and SLCL were similarly influential with regard to learners' decisions in the vocabulary test. The aim of Experiment 5 was to examine the nature of both mechanisms by studying the interaction of CSWL and SLCL when both are independently applicable: As in Experiment 4, nouns had two potential meanings; this time, however, both CSWL and SLCL supported the high-frequency meaning. Results clearly provide evidence for the hypothesis that SLCL completely blocks learner's sensitivity to smaller difference in co-occurrence frequencies which characterizes pure CSWL learning. In contrast, SLCL increased learners' sensitivity to category membership. This pattern confirms the hypothesis that while CSWL is a parallel, probabilistic, and incremental way of learning, SLCL is more deterministic and semantically based. Additionally, results from a vocabulary test one day after learning reveal that learning rates were still clearly above chance.

Taken together, our experimental data clearly shows that CSWL and SLCL are powerful mechanisms for word learning in adults in non-instructed environments, which may lead into long-lasting retention. Importantly, these mechanisms interact in multiple ways due to differences in their nature: They can be used in a complementary way (either to learn a set of nouns, or to learn a single noun meaning), they influence word learning about equally strongly when they are in conflict, and SLCL blocks CSWL when both mechanisms are independently (i.e., redundantly) applicable. We conclude that adult word learners follow an efficient strategy: They employ as many resources in parallel as necessary but ignore the less direct and helpful cue when information is redundant. However, when the relevance of different cues is unclear, they consider all of them.

Zusammenfassung

Eine zweite Sprache zu erlernen, ist eine schwierige Aufgabe für Erwachsene. Das Verstehen und Lernen unbekannter Wörter ist mühsam, insbesondere in nicht-instruierten Situationen: Wörter sind oft Teil komplexer linguistischer Kontexte und potentielle Referenten sind eingebettet in visuelle Szenen, die keine eindeutigen Zuordnungen erlauben. Um diese Schwierigkeiten zu überwinden, haben Lerner jedoch die Möglichkeit, die Reichhaltigkeit ihrer multi-modalen Umgebung zu ihrem Vorteil zu nutzen, indem sie eine Reihe verschiedener Wortlern-Mechanismen anwenden. Automatische Sprach-Verarbeitungs-Mechanismen unterstützen die zügige Integration unterschiedlicher Informationen.

Obwohl viele Wissenschaftler aus diversen Bereichen Studien zum Wortlernen durchgeführt haben, gibt es nur sehr wenige Untersuchungen zum Zusammenspiel verschiedener Wortlern- und Verarbeitungs-Mechanismen. Dies jedoch führt möglicherweise dazu, dass das Szenario gleichzeitig als zu einfach und als zu kompliziert dargestellt wird. Überdies sind die Versuchsstimuli der meisten Studien entweder durch fehlende Natürlichkeit oder durch fehlende Kontrolle gekennzeichnet.

Das Hauptziel dieser Arbeit ist es, Wortlernen bei Erwachsenen in einer stärker integrativen und interaktiven Weise zu erforschen. Hierbei sollen verschiedene Informationsquellen sowie Wortlern-Mechanismen und -Prozesse gleichzeitig berücksichtigt werden. Die Motivation für diese Auseinandersetzung ist es zu einer vollständigeren Theorie des Wortlernens in einer Zweitsprache beizutragen, Forschungstraditionen zu verbinden und Rückschlüsse für die Entwicklung praktischer Lern-Anwendungen zu ziehen.

Im Speziellen haben wir die Interaktion zweier bedeutender und visuell integrierter Wortlern-Mechanismen untersucht, Situations-Übergreifendes Wortlernen (*cross-situational word learning*, CSWL, Yu and Smith, 2007) und Wortlernen, das auf Beschränkungen bezüglich der Satzebene basiert (*sentence-level constraint learning*, SLCL). CSWL funktioniert bottom-up: Lerner ziehen Verbindungen zwischen visuellen Objekten und gesprochenen Wörtern, indem sie die Häufigkeiten ihres gemeinsamen Auftretens (Mit-Auftretens) verfolgen. SLCL vollzieht sich im Gegensatz dazu top-down, denn es folgt dem Grundsatz des Inferierens basierend auf dem linguistischen Kontext und gegebenenfalls zusätzlichen Informationen. SLCL in dieser Arbeit bezieht sich auf das Inferieren von Objekt-Nomina-Bedeutungen (z.B. *der Maiskolben*) auf Grund restriktiver Verben (z.B. *essen*), einem visuellen Kontext, und dem Weltwissen des Lerners.

Der Kern unserer Studien ist ein neuartiges experimentelles Paradigma, in dessen Rahmen erwachsenen Deutschen Muttersprachlern in einer stufenweisen Prozedur eine semi-natürliche Mini-Sprache gelehrt wird. Partizipanten wurden zunächst mit einer Reihe Verben vertraut gemacht (z.B. *bermamema* 'essen'), bevor ihnen Materialien zum Nomina-Lernen dargeboten wurden. Die experimentellen Items bestanden aus visuellen Szenen, die mit gesprochenen transitiven Sätzen gepaart waren (z.B. *Si laki bermamema si sonis.* 'Der Mann isst den Maiskolben'). Die zu lernenden Nomina waren also Teil von Sätzen. Am Ende des Experiments wurde ein selektiver Vokabeltest (mit Konfidenz-Selbst-Wertung) durchgeführt. Während des Lernens und Testens wurden die Augenbewegungen der Versuchsteilnehmer aufgezeichnet.

In Experiment 1 haben wir den Gebrauch von CSWL und SLCL in diesem Paradigma evaluiert, das heißt, in einer Situation, die natürlicher ist als in bisherigen Wortlern-Experimenten. Unsere Ergebnisse zeigen, dass die Teilnehmer beide Mechanismen komplementär angewendet haben, um die Vokabeln zu lernen. Zudem bewirkte SLCL eine klare Verbesserung der On-line Identifizierung von Referenten. In Experiment 2 wurde eine zweite Wortreihenfolge eingeführt (OVS), die dadurch charakterisiert ist, dass das Verb dem syntaktischen Objekt (welches ein Objekt in der Szene bezeichnet) folgt statt ihm voraus zu gehen (wie in SVO). Die gewonnen Resultate stimmen mit der Hypothese überein, dass die Verb-basierte Vorhersage von Referenten einen positiven Einfluss auf das Erlernen von Nomina hat. In Experiment 3 gingen wir erneut der Frage nach, ob SLCL Nomina-Lernen verstärkt und wendeten uns der Interaktion von CSWL und SLCL zu, indem wir den Grad der Ambiguität von Nomina-Referenten-Zuordnungen manipulierten. On-line- und Off-line-Ergebnisse liefern Evidenz für die Hypothesen, dass SLCL das Erlernen von Nomina verstärkt und, dass CSWL und SLCL interagieren, indem sie gemeinsam zu der Identifizierung von Nomina-Bedeutungen verhelfen. Experiment 4 wurde durchgeführt um die Interaktion von CSWL und SLCL zu erforschen, wenn beide Mechanismen konfligieren. Experimentelle Nomina hatten zwei potentielle Bedeutungen mit unterschiedlichen Mit-Auftretens-Wahrscheinlichkeiten. Da die verbalen Restriktionen den Referenten mit der geringeren Mit-Auftretens-Wahrscheinlichkeit unterstützte, standen CSWL und SLCL in Konflikt. Gewonnene Lernraten zeigen eindeutig, dass CSWL und SLCL einen gleichermaßen starken Einfluss auf die Entscheidungen im Vokabeltest hatten; Augenbewegungen jedoch deuten an, dass CSWL auf der Aufmerksamkeits-Ebene leicht dominierte. Das Ziel von Experiment 5 war es, der Beschaffenheit von CSWL und SLCL auf den Grund zu gehen, indem ihre Interaktion in einer Situation erforscht wurde, in der beide Mechanismen unabhängig voneinander anwendbar waren: Wie in Experiment 4 hatten alle Nomina zwei Bedeutungen, jedoch unterstützten dieses Mal sowohl CSWL als auch SLCL den Referenten mit der höheren Mit-Auftretens-Wahrscheninlichkeit. Experimentelle Ergebnisse belegen, dass SLCL die Sensibilität für kleinere Unterschiede in der Mit-Auftretens-Wahrscheninlichkeit blockiert, durch welche reines CSWL-Lernen gekennzeichnet ist. Stattdessen erhöhte SLCL aber die Sensibilität für Kategorie-Zugehörigkeiten. Diese Muster bestätigt die Hypothesen, dass CSWL parallel, probabilistisch und inkrementell verläuft, während sich SLCL eher deterministisch und kategorien-basiert vollzieht. Die Resultate von einem Vokabeltest einen Tag nach der Lernprozedur zeigen zusätzlich auf, dass die Lernraten noch immer deutlich besser waren, als es zufällig zu erwarten wäre.

Zusammengenommen belegen unsere experimentellen Daten eindeutig, dass CSWL und SLCL wirkungsstarke Mechanismen für das Wortlernen bei Erwachsenen in nicht-instruierten Kontexten sind, welche wahrscheinlich langfristige Lernerfolge nach sich ziehen. Auf Grund ihrer unterschiedlichen Beschaffenheit interagieren diese Mechanismen in vielerlei Hinsichten: Sie können komplementär angewendet werden (entweder um eine Gruppe von Woȓter oder um ein einziges Wort zu lernen), ihr Einfluss auf das Erlernen eines Wortes ist ungefähr gleich groß, wenn sie konfligieren und SLCL blockiert CSWL in Fällen, wenn beide Mechanismen unabhängig voneinander (d.h. redundant) angewendet werden können. Wir ziehen die Schlussfolgerung, dass erwachsene Wortlerner einer effizienten Strategie folgen: Sie machen von so vielen Ressourcen wie notwendig parallel Gebrauch, ignorieren aber weniger direkte und hilfreiche Hinweise, wenn redundante Informationen zur Verfügung stehen. Ist die Relevanz verschiedener Hinweise jedoch unklar, so berücksichtigen sie jede dieser Informationsquellen.

Acknowledgements

I was looking forward to writing these lines for some time already. This is, apparently, partly because this is one of the very last bits that have to be done before I will get rid of this beloved little monster called thesis but also because I think that it is actually a really great thing to sometimes say thank you to people who do thankworthy things. It is exactly these people who made writing this thesis not only practically possible but also a great time despite of all hassle.

My first big thank you goes to Matthew Crocker for supervising and supporting me in a very friendly and cooperative manner and for, unlike myself, never loosing patience even when things did not go the planned way. I also thank you for providing us with perfect equipment and facilities which make conducting research a great deal more comfortable.

I am very thankful to my charming colleagues Maria Staudte, Emilia Ellsiepen, Afra Alishahi, Heiner Drenhaus, Pirita Pyykkönen, Juliane Steinberg, and Berry Claus for interesting discussions, useful suggestions, and necessary assistance with computer programs (thank God you guys are so smart...). I would more than that like to thank you and other fun people in Saarbrücken for many happy hours inside and outside University without which this time would not have been a tenth as enjoyable as it was. Special thanks to lovely Heiner and my highly estimated friends Emilia and Maria for expert proof reading and cutting moreovers and firstlies. Many thanks to the German Research Foundation (DFG) for financially supporting my research in Saarbrücken as well as within numerous exciting conference journeys. I would also like to thank Martin Pickering (and people in Edinburgh) for making me a psycho- and Nikolaus Himmelmann (and people in Bochum) for making me a linguist.

I am most honestly grateful to my parents and brother for unconditional trust and support (and for never asking me why I do not study something useful). Moreover, I want to say thank you to my friends Jana Wiechers and Sonja Riesberg for 'open ears and arms' at any time. Finally, a very special thank you to André Fuetterer for accompanying and enriching my life.

Contents

List of Figures

List of Tables

1. Second-Language Word Learning in Non-Instructed Contexts - Challenge and Opportunity

Learning a second language is a demanding undertaking for adults. In an instructed environment like a classroom, language input and required output are traditionally organized in followable steps and controlled by a supervisor. In more realistic settings outside classroom, however, the task becomes more challenging: Spoken language is complex, embedded in rich communicative situations, and situated within linguistic and visual contexts. Moreover, language input needs to be not only comprehended quickly and often reacted to immediately but it also needs to be exploited as continuing learning source (Field, 2007).

Word learning, essential for gaining the command of a language, becomes intriguing: Firstly, the sentential context that novel words occur in is often rich. That means that understanding a particular word is inter-dependent with understanding other words as well as grammatical structures. Secondly, the visual context is *noisy* and its contribution is uncertain: It may or may not contain relevant information such as the referent a novel word denotes; if it does contain this referent, it is probably not easily identifiable because many other entities and relations are in the scene concurrently, which means that the referent-word relation (or *world-word relation*) is still ambiguous (Quine, 1960). In addition, the visual environment is dynamic in that it undergoes constant change.

However, every challenge presents an opportunity: While people constantly have to deal with their rich environments, they are very good at also exploiting it. They can precisely keep track of useful information coming from all available modalities such as various linguistic levels, the visual scene, and experiences with the world (e.g., Kirkham, Slemmer & Johnson, 2002). Human sentence-processing mechanisms are known to integrate this *multi-modal* information automatically and *incrementally* (i.e., in a nutshell: rapidly and step-wise) for language comprehension (Tanenhaus & Trueswell, 1995). Importantly, this also helps language novices to find the meanings of novel words: A variety of different word-learning mechanisms rely on the integration of multi-modal cues (e.g., Yu & Smith, 2007; Landau & Gleitman, 1985). Some of these mechanisms are more important for vocabulary learning in *first language acquisition* (FLA), whereas others are more (or only) relevant to *second language acquisition* (SLA); some mechanisms are used more *intentionally* (i.e. with the explicit intention to learn), others more *incidentally* (i.e. without the explicit intention to learn). While FLA word learning has exclusively been investigated as non-instructed, SLA word learning in non-instructed environments is highly under-researched. This may have caused a tacit under-estimation of the relevance of some learning mechanisms for SLA word learning. Furthermore, while there are many studies investigating the different word-learning mechanisms, their interplay is surprisingly under-researched (both for SLA and FLA), which constitutes a notable gap for a complete comprehension of word learning: One the one hand, word learning may be easier when multiple mechanisms are available; one the other hand, it is unclear how different mechanisms influence one another and whether this may complicate the situation.

Imagine you are a learner of English. You already understand most of the grammar and many words but you still encounter knowledge gaps from time to time. One day you are drinking a mixed fruit juice. On the label of the bottle you find a picture of all contained fruits as well as a list of ingredients. You see on the picture that there are bananas, apples, oranges, kiwis, mangos, pears, and two fruits that you do not know (a red one and a green one). Now you check the list of ingredients: 'apple juice (35%), orange juice (31%), kiwi nectar (12%), mango nectar (6%), banana nectar (5%), dax nectar (4%), pear nectar (4%), gorp concentrate (3%). Given that you neither know how gorps look like nor how daxes look like, either of these two nouns could be the name of either of the two unknown fruits. That means that you can only infer that either 'dax' is the name of the red fruit and 'gorp' is the name of the green fruit or 'dax' refers to the green one and 'gorp' to the red one. Two days later you are on the market and you observe how a woman is buying a flower that you cannot name and two sorts of exotic fruits - a yellow fruit and the red one which was also in the juice. The woman is asking the vendor 'Do I have to peel the dax before eating?'. Given that people do not tend to eat flowers and that this is the second time already that you encounter the name 'dax' together with this red fruit (while the yellow has co-occurred with 'dax' only once), you can now make the hypothesis that 'dax' denotes the red fruit.

Researchers from various fields are interested in word learning:[1] child acquisition researchers, second language researchers, theoretical linguists and philosophers, psycholinguists and applied linguists, language teachers, language assessment associates, software developers, as well as speech therapists and neurologists. These groups, however, have diverging objectives and, importantly, they apply different methods, all of which have their strengths and limitations. One trade-off regarding the methodologies used to examine word learning

[1]From now on I will use the terms *learning* and *acquisition* in an interchangeable way which is neutral to any discussion about innateness, maturational constraints, intentionality, or consciousness.

is between control and naturalness. Psycholinguists and child language researchers, for instance, tend to use well controlled procedures which, on the downside, often suffer from a high degree of unnaturalness: Firstly, individual word-learning mechanisms are rarely studied in a combined and interacting way. Secondly, the input is highly isolated and simplified, in particular, words are usually not presented as parts of sentences and visual referents are not embedded in coherent scenes. This over-complicates the scenario because not all sources of information that may crucially contribute to learning are available while at the same time over-simplifying the situation because the learning environment is less noisy and complex than in reality. Studies by second language researchers and language teachers, in contrast, tend to be more naturalistic but often rather uncontrolled. The drawback here is that it is difficult to draw clear and precise conclusions about the factors which are at play.

This thesis investigates word learning in an integral manner, considering both various available information sources and different learning and processing mechanisms and processes. We will argue that this is necessary to approach a complete and realistic theory of second language lexical acquisition because second-language word learning is an inherently *interactive* and *constraint-based* process that cannot be fully understood if different factors are not considered simultaneously. We moreover attempt to study these mechanisms in a more natural way than it has been done in most other laboratory studies, while still maintaining careful control, in order to draw clear conclusions which are relevant for natural learning settings. By doing so, we additionally follow the aim to bridge research fields on word learning. A final motivation driving this investigation is to examine learning environments from a more practical perspective, in order to draw conclusions for applications.

The nature and interplay of two word-learning mechanisms will be focussed on in this thesis: *cross-situational word learning* (**CSWL**, Yu & Smith, 2007) and *word learning based on sentence-level constraints* (*sentence-level constraint learning*, **SLCL**). To learn words cross-situationally, people must keep track of the *co-occurrence statistics* of word-referent combinations across situations. That is, they have to monitor which words and referents occur together. SLCL, on the other hand, exploits systematic syntactic and semantic relations between sentence parts (words and constituents) to constrain hypotheses about novel word meanings (e.g., *syntactic bootstrapping*, Landau & Gleitman, 1985). Specifically, we are investigating word learning via semantic constraints that *restrictive verbs* (such as *eat*) impose on their direct objects, together with the visual context (e.g., frisbee) and people's world knowledge (see Altmann & Kamide, 1998). We will motivate the selections of these two particular mechanisms in Chapter 2.

In the remainder of this thesis, we will first begin by reviewing a range of findings on, firstly, word-learning mechanisms in FLA and SLA and, secondly, human sentence-processing mechanisms in people's first language (L1) and people's second language (L2) (Chapter 2). This theoretical overview will identify notable gaps existing in research and motivate the way we experimentally investigate word learning. Five psycholinguistic experiments will then be presented, addressing the two learning mechanisms CSWL and SLCL as complementary (Chapter 3), conflicting (Chapter 4), and independently applicable (Chapter 5), in order to weigh their impact, expose their exact interplay, and, finally, examine their underlying nature. At the conclusion of this thesis, results and implications as well as possible applications will be discussed (Chapter 6).

2. Word-Learning and Sentence-Processing Mechanisms

2.1. Word Learning

Learning the vocabulary of a language is a central aspect of both first language acquisition (FLA) and second language acquisition (SLA; Read, 2004; Davis, Di Betta, Macdonald, & Gaskell, 2008; Richards, 2000; Tight, 2010). Since the meaning of a sentence is fundamentally built from the meanings of the words contained in it, vocabulary knowledge is essential for comprehension (Richards, 2000), whereas it is possible to understand the rough sense of a discourse without being proficient in grammar (Schmitt, 2000; Barcroft, 2004). At least 1000 words must be known productively before language novices can express themselves and the receptive command of even around 3000-5000 vocables is necessary before most texts or utterances can be (close to completely) understood (Laufer, 1992: 3000 words; Nation & Waring, 1997: 3000-5000 words; Hulstijn, 2001: 5000 words).

Words, in particular names of concrete objects and actions, are the first linguistic knowledge that infants acquire and which enable them to start communicating (Gillette, Gleitman, Gleitman, and Lederer, 1999). This is not accidental: Children experience the world and have sensorimotor mental representations for concepts before they begin talking (Howell, Jankowicz & Becker, 2005). The meanings of concrete content words is the part of language which is most grounded in these experiences, or in other words, these words are the easiest and most direct way to start connecting the world to the arbitrary system of language (Lakoff, 1987).

The first two questions that have to be clarified when talking about word learning are, on the one hand, what a word is, and on the other hand, what it means to know a word's meaning (Read, 2000; Gass, 1988; Snedeker, 2008). The construct 'word' can be defined linguistically as a language's smallest meaning-bearing unit (Read, 2000). It is a matter of debate what exactly counts as a word but we leave this issue aside. In this thesis, the focus will be on learning the meanings of concrete nouns only, which clearly belong to the class of words.

The second question about what it means to have learned a word is more relevant for this thesis. It is an important distinction whether a learner has only heard a word before, roughly understands it, or knows what it means even in different contexts and correctly uses it productively (Read, 2000). Nation (1990: p.31) suggests that knowing a word includes mastering its meaning(s), the written form, the spoken form, the grammatical behavior, possible collocations, the register, as well as the associations and the frequency of the word. Knowing a word in all its facets can therefore only be an incremental process which requires many exposures (Schmitt, 2000). Interestingly, recent research has shown on the contrary that novel words can be relatively deeply integrated into the mental lexicon after a very short time already: Priming and fMRI results reveal that novel words participate in lexical competition after 12 hours, including a sleep period (Lindsay & Gaskell, 2010 for a summary). Moreover, an MEG study by Dobel, Junghfer, Breitenstein, Klauke, Knecht, Pantev & Zwitserlood (2010), provided evidence for the hypothesis that words which were learned (via statistical training) over five days, were integrated into conceptual and lexical memory networks: The newly acquired words primed corresponding pictures in a similar manner as do L1 words. This, of course, does not mean that the word is perfectly known in all the ways listed by Nation. Usually, the receptive vocabulary knowledge precedes the productive command and learners' (L1 and L2) receptive vocabulary is always larger than their productive repertoire (L1: Goldin-Meadow, Seligman & Gelman, 1976; L2: Melka Teichroew, 1982; Laufer & Paribakht, 1998; Webb, 2008).

In this thesis, word learning is mainly assessed at an initial stage: The primary measurement is associative testing which takes place after about half an hour of exposure to the novel language. However, Experiment 1 includes a sentence-judgement test whose results highly correlate with those of the forced-choice noun test, suggesting a deeper level of acquisition than mere superficial association. Experiment 5 also includes a vocabulary test one day after learning which shows only little change in performance. Thus, while we base our claims on the first stages of word learning, we nonetheless expect that the learning mechanisms under investigation contribute to long-term learning on a deeper level (see also discussion about cross-situational word learning below).

When reviewing studies on first and second language acquisition in general and on word learning in particular, it is conspicuous that these two domains have relatively different foci. Firstly, while children's task is to find out how the world they experience is mapped onto language, adult learners already understand this systematic relationship and most concepts to be labelled in the target language already have a name in their first language (MacWhinney, 2005). Secondly, first-language word learning is rather observed as automatic and researchers attempt to understand this mysterious ability, whereas the aim in second language acquisition is primarily to examine, develop, and teach strategies, and research is therefore more goal-oriented. This difference is also reflected in the terminology that researchers use: Children are learning *words* and adults are aiming at gaining the command of a language's *vocabulary*.

Word acquisition in infants is fascinating because it happens reliably and appears to be effortless. Successful lexical acquisition nevertheless requires very complex (and human specific) cognitive behavior. First of all, the child needs to segment the speech stream to identify a word's unique sound. Secondly, a link must be established between this acoustic pattern and some meaning or function. To find the meanings of unknown words, children naturally make use of a range of unconscious mechanisms: They attend to and observe the visual world in which language and they are situated (Carey, 1978), they are sensitive to statistical regularities (Smith & Yu, 2008) and social cues (Baldwin, 1993; Tomasello, 2000), and they exploit linguistic contexts (Landau & Gleitman, 1985). Importantly, they are also guided by innate biases and constraints (Markman, 1990). These mechanisms will be elaborated on in more detail in Section 2.2.

Word learning (or vocabulary/lexical learning) in second language acquisition, on the other hand, is a much more conscious and goal-oriented process and seems to be more effortful (Read, 2004). However, it is not the case that adults are bad word learners: In contrast with other aspects of language acquisition, adults are at least as good as children (Snedeker, 2008). While children exclusively learn words *incidentally* or *implicitly* (Schmitt, 2000), adults can additionally make use of *intentional* (or *explicit*) learning strategies. Incidental learning means that somebody learns something without intention and instruction, for instance as in picking up novel words while reading (Barcroft, 2004). Importantly, this does not mean that learning is unconscious. Intentional learning, on the contrary, is goal-oriented learning (e.g., vocabulary learning via lists). As Barcroft (2004), points out, both concepts are best considered as poles of a continuum rather than a dichotomous classification. MacWhinney (2005; 2008) notes that although L1 learning mechanisms are less powerful in L2 than L1, they are still very relevant for both. He considers L1 learning mechanisms as a subset of L2 learning mechanisms:

> '... [T]he method for learning new word forms in a second language is basically an extension of the methods we used for learning words in our first language.' (MacWhinney, 2004: 2)

2.2. Word-Learning Mechanisms in FLA and SLA

The mechanisms at work in child- and adult-word learning will now be considered in greater detail. In Sections 2.2.1 to 2.2.4, mechanisms which are relevant in FLA are discussed. For some of them, there is also clear evidence for their relevance in SLA. However, while all of these mechanisms are incidental for FLA (as mentioned above), this is not always the case

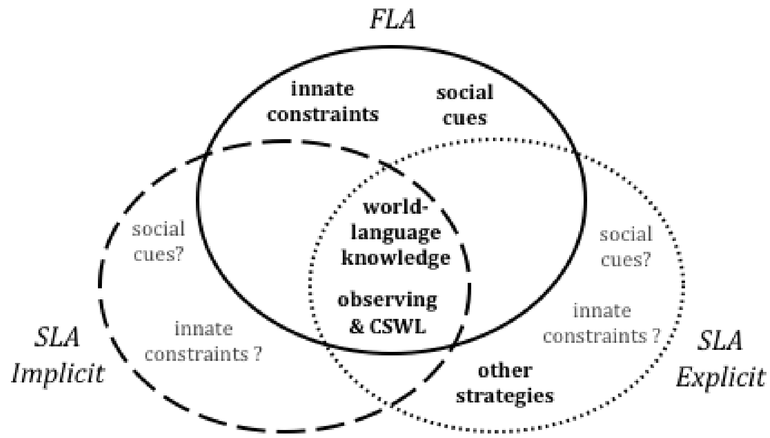


Figure 2.1.: Word-learning mechanisms

for SLA: Some processes which are unconsciously (and therefore also non-intentionally) influencing children are (also or only) useful as explicit strategies in adults. That means that the classes of word-learning mechanisms belonging to FLA and SLA and implicit and explicit are all partially overlapping (see Figure 2.1). Section 2.2.5 discusses implicit and explicit word-learning mechanisms that are only relevant for SLA. In section 2.2.6, the interplay of word-learning mechanisms will be addressed.

2.2.1. Observational Cues and Cross-Situational Word Learning

Observation is generally a promising activity for a language novice because the people, objects, actions, properties, and relations which novel words describe are grounded in the world. While the world is most easily connected to concrete nouns, the acquisition of verbs and adjectives can also be facilitated (e.g., Childers & Paik, 2008). Given that there is exactly one unknown word (in a known linguistic context) and a clear relationship with exactly one object (for instance because it is the only object in view), infants can *fast-map* this word onto the object (Carey, 1978), that means, they can build lexical representations of that word after only one exposure. This appears to be relatively easy but realistic scenarios tend to be more challenging: The visual environment is usually noisy, that is, potential relationships between bits of language and things in the world are highly ambiguous. This noise causes two basic problems: The first is to identify what a speaker is talking about (*reference problem*), that is, to select a visual target. The second problem, referred to as *frame problem* (as applied to word learning; Fodor, 1987; Hayes, 1987; Nappa, Wessel, McEldoon, Gleitman & Trueswell, 2009) is to understand which concept the speaker exactly refers to (e.g., which level of specificity). As Quine (1960), illustrated (and hundreds of researchers have investigated), the mere presence of, for example, a rabbit, together with an unknown word (*gavagai*) open up a vast number of possible mappings such as *rabbit*, *animal*, *rabbit's nose*, *sitting*, or *white and fluffy*, just to name a few. Both the reference problem and the frame problem are sometimes described under the labels *referential uncertainty* (Gleitman, 1990; Vouloumanos & Werker, 2009) or *indeterminancy problem* (Yu & Smith, 2007), and are relevant for both children and adults.

One way to overcome these problems is to track statistical regularities between words and instances, relations, and properties in the world over time (Quine, 1960). This analysis of co-occurrences of words and objects (or relations or properties) has been referred to as *statistical word learning* or *cross-situational word learning* (**CSWL**; Quine, 1960; Siskind, 1996; Akhtar & Montague, 1999; Vogt & Smith, 2005; Yu & Smith, 2007). Figure 2.2 illustrates a possible CSWL setting: A learner of English encounters two situations with two spoken words each (Utterance 1: *ball* and *bat*, Utterance 2: *dog* and *ball*) and two visual referents each (Scene 1: ball and bat, Scene 2: ball and dog). Given only the first situation, the learner cannot differentiate whether the ball is referred to as *ball* and the bat is referred to as *bat* or, alternatively, if the name of the ball is *bat* and the name of the bat is *ball*. However, having seen both the first and the second situation, he will make the hypothesis that the referent for the noun *ball* is the ball because both scenes contain a ball and both utterances contain the word *ball* whereas there is no further overlap between the situations. That means that, given these two situations, the probability (P) of seeing a ball when the word *ball* is spoken is 1.

Smith & Yu (2008) found that 12 and 14-month-old children are successful and quick in tracking the co-presence of isolated novel nouns and unknown (also isolated) objects across trials when referential uncertainty in one trial was 2:2 (per trial, two objects and two spoken words were presented), as in the example explained above. Numerous other CSWL studies have been conducted (most of them over the last 6 years), with both children and adults, mainly with nouns but also with other parts of speech, with implicit and explicit instructions, with different levels of complexity, with high and low referential uncertainty, and in various languages (e.g., Siskind, 1996; Akhtar & Montague, 1999; Vogt & Smith, 2005; Yu & Smith, 2007; Smith & Yu, 2008; Vouloumanos, 2008; Vouloumanos & Werker, 2009; Childers & Paik, 2008; Monaghan & Mattock, 2009; Kachergis, Yu & Shiffrin, 2010b; Breitenstein & Knecht, 2003). Importantly, Kachergis et al. (2010b) found that while CSWL in adults is successful both with and without instruction (i.e., incidentally), it is superior in the condition including explicit instruction.

Recently, there have been investigations going further into depth regarding the exact mechanisms operating in CSWL and its underlying nature: Success in CSWL has been shown to depend on different factors such as the frequency of exposure of referent-word

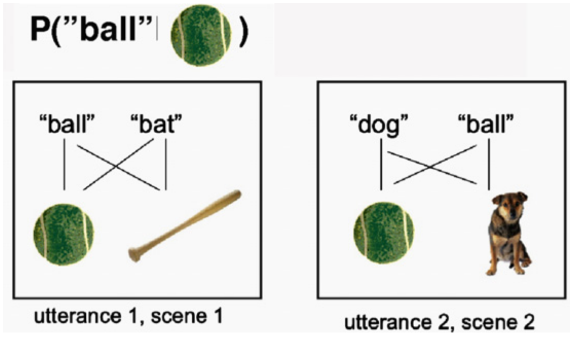


Figure 2.2.: Cross-situational word learning (example from Smith and Yu, 2008: 1560)

pairs, the diversity of other referent-word pairs and the degree of within-trial ambiguity (Kachergis, Yu & Shiffrin, 2009a), the distance between single exposures (Kachergis, Yu & Shiffrin, 2009b), the conditional probability of a word given a referent and vice versa as well as the joint probability of both occurring together (Klein & Yu, 2009) and the order of high-informative and low-informative trials (Medina, Hafri, Trueswell & Gleitman, 2010).

The precise manner in which information accumulates across trials in CSWL has also been a recent matter of discussion. A common assumption is that CSWL is an incremental adding up of evidence which implies that all collected information is available any time (Smith, Smith & Blythe, 2010). That means that when an initial hypothesis about a world-word mapping must be rejected because the object which was assumed to be the referent is not co-present with the word in the current scene, the learner will switch to hypothesizing that the correct referent is exactly that object in the scene which has the highest co-occurrence frequency with the word across all trials. Smith and colleagues (2010) refer to this strategy as *pure cross-situational word learning*. The authors argue that there are also other ways word learners can learn cross-situationally. In contrast with pure CSWL, they identify the strategy of *approximate cross-situational word learning*: When a hypothesis must be changed, learners do not consider the other co-occurrence frequencies across all situations but randomly pick a novel assumed referent from the current scene. Smith et al. present experimental evidence suggesting that both of these two strategies are applied, depending on the complexity of the task: When the learning scenario is simple enough, pure CSWL is conducted, otherwise learners use approximate CSWL.

An opposing view comes from Medina et al. (2010) who argue that CSWL works exclusively via fast-mapping-like situations in the beginning of learning, followed by situations in which this hypothesis can be confirmed. However, recent data from Vouloumanos (2008), Vouloumanos & Werker (2009), and Yurovsky, Fricker, Yu & Smith (2010) supports the hypothesis that CSWL is (or at least can be) an incremental and parallel way of learning.

This issue will be further discussed in Chapter 5.

It is only a first step for word acquisition to identify a referent (*labeling*, Aitchinson, 1987). Next, the word must become more deeply integrated into the mental semantic network and all ways that it can be used must be understood (*slow mapping*, Carey, 1978; *categorization* and *network building*, Aitchinson, 1987). Nation 1990 emphasizes the importance of repeated exposure and even makes the estimate that it requires five to sixteen repetitions before a word is acquired (see also Barcroft, 2004). Given that CSWL is based on repetition, it is therefore a potentially realistic way of word learning, which can lead into deep word knowledge (Blythe, Smith, and Smith, 2010). The MEG study by Dobel, Junghöfer, Breitenstein, Klauke, Knecht, Pantev & Zwitserlood (2010), mentioned above, in fact revealed that words which were just learned in a statistical way were indeed integrated into memory (see also Breitenstein, Zwitserlood, De Vries, Feldhues, Knecht & Dobel, 2007 for a similar study without MEG). In their experiment, German participants were trained on 90 novel spoken words. Per trial, one word and one line drawing were presented. Each word was repeated 40 times in the course of five consecutive days (each day maximally 20 minutes). 45 words belonged to the learned set and 45 words to the unlearned (control) set: The learned-set words were presented 20 times with the target picture and two times each with ten other pictures. Control words were paired arbitrarily. Participants first heard the word and saw the picture and were asked to decide whether picture and word matched or not. Results revealed that on Day 5, 93% of the words from the learned set were learned (Day 1: 53%). Additionally, as described above, MEG signals, which were recorded during cross-modal priming tests conducted before and after training, revealed that training caused the novel words to be integrated into memory: When used as prime for the trained-on depiction, there was a reduced mismatch effect in the post-test compared to the pre-test (reduced N400m). Brain responses even resembled those elicited by native words.

The CSWL studies cited above build a very solid base from which to argue that CSWL is indeed a potentially powerful mechanism in language learning, among others for adults in non-instructed environments. However, almost all studies suffer from a high degree of unnaturalness: Words are neither embedded in linguistic contexts nor are objects parts of natural coherent visual environments. This on the one hand oversimplifies the situation, because it eliminates complexity, but it also potentially overcomplicates the scenario because the linguistic and visual context can contain helpful information. A central aim of this thesis is to address these issues.

2.2.2. Social Cues

Another way the visual context can be made less ambiguous and therefore more helpful for word learning children is people's social behavior: By studying natural mother-child interactions, Tomasello and colleagues found that caretakers often directly attend to objects that they are talking about, they hold them, point at them, or look at them (e.g., Tomasello & Todd, 1983). Adjusted intonation can also be assisting (Houston-Price, Plunkett & Duffy, 2006). Children have been shown to be very sensitive to these social cues and they use them to facilitate word learning (Baldwin, 1991; Baldwin, Markman, Bill, Desjardins, Irwin & Tidball, 1996; Tomasello & Barton, 1994; Carpenter, Nagell, Tomasello, Butterworth & Moore 1998; Baldwin 2000; Butterworth, 2004; Nappa et al., 2009). It is a matter of ongoing discussion, when children develop which grade of sensibility for social cues. Most likely, they first have to overcome a stage of egocentricity in the first months of life before they are able to interpret gaze and other cues in an advanced and intention-reading way (see Shepherd, 2010).

Little research investigating the role of social cues for adult language learners has been conducted. While it is relatively clear that they are less available for adults because infants receive more social support (Snow, 1999; MacWhinney, 2008), gaze and gesture are still potentially beneficial for second language learners in conversations. There is some evidence for the positive role of social cues in second language learning: Allen (1995) demonstrated that students performed better in vocabulary retention of french expressions when they were presented gestures during training (e.g., *Je m'enlave les mains*, 'I wash my hands of the whole affair.' together with a hand-rubbing gesture). However, since learners also repeated the gestures themselves, it is unclear what exactly caused the effect (observing gestures or enacting gestures). Sueyoshi & Hardison (2005) presented interesting data which reveal that the comprehension of lower proficiency learners was better with natural gestures performed by the speaker (iconic, metaphoric, beat, and deictic) than without. Surprisingly, however, comprehension in higher proficiency learners was better without gestures. While gestures have additionally often been claimed to be a means of highlighting an unknown word in order to guide learners' attention (*input enhancement*, Sharwood Smith, 1991; Sharwood Smith, 1993; Schmitt, 1997), there is no direct experimental evidence for this hypothesis. Given that adults integrate gestures jointly with speech into their comprehension (Kendon, 2004; McNeill, 1992) and that gestures can facilitate native language comprehension (Kelly, Barr, Breckenridge Church & Lynch, 1999), it is plausible to assume that they can also assist the comprehension of second language input (implicitly or explicitly or both). Deictic gestures (pointing) can directly identify referents for novel nouns and iconic gestures (e.g., drawing a shape with the hands) potentially provide partial meaning. While social cues are potentially interesting regarding adult word learning, this topic will not be considered in this thesis.

2.2.3. Word Learning Based on Sentence-Level Constraints

Words can moreover be learned via sentence-level constraints (from now on referred to as *sentence-level constraint learning*, **SLCL**). Given that the learner has some prior knowledge about the language's grammar (e.g., about word order) and vocabulary, the linguistic context in which a novel word occurs can be very informative. In general, the more a learner already knows about a sentence's structure and the words contained in it, the easier it is to understand the meaning of an aspect which is missing (Christophe, Millotte, Bernal & Lidz, 2008). This is partly because all languages have systematic grammatical regularities, for instance concerning both the relations between parts of a sentence and the distributions of words belonging to certain categories. The linguistic context can also interact with knowledge that people have about the world, eliciting expectations about the semantic content of a word. Finally, both cues, the linguistic context and world knowledge, often convey information together with the visual world. In the following, we will first review research on word learning via the linguistic context which investigates precise effects of constrained linguistic contexts such as verb frames (*bootstrapping effects*). Second, the more general role of the linguistic context with regard to SLA lexical learning will be examined.

2.2.3.1. Linguistic Bootstrapping

One transparent linguistic relation which can facilitate word learning is the one between verbs and their arguments. A transitive causative verb such as *to hug*, for instance, requires two arguments, a subject (a noun referring to an agent, the hugger) and a direct object (a noun referring to a patient, the hugged). When people encounter this verb frame with a novel verb (e.g., *The boy daxes the girl.*), they can infer that the verb *to dax* is likely to have a causative meaning (*syntactic bootstrapping*, Landau & Gleitman, 1985). If there is additionally a scene depicting a causative event (e.g., hugging), even more precise hypotheses about the verb's meaning can be made. In fact, many theorists argue that verb learning in infants is only possible via this process (e.g., Gleitman & Gleitman, 1992). There exist a number of studies revealing that children exploit the systematic syntax-semantics relations of argument structure to infer verb meanings (e.g., Brain, 1992; Fisher & Gleitman, 2002; Fisher, 2002; Lidz, Gleitman & Gleitman,2003; Naigles & Swensen, 2007). Lee & Naigles (2008), for instance, showed that 2- and 3-year-old Chinese children use the causative subcategorization frame (someone VERBs someone else) to infer causative meanings of novel verbs. As Arunachalam & Waxman (2010b) reported, importantly, it is really the syntactic frame itself which is important for the bootstrapping effect: Providing loose information about referents only (*Let's see a boy and a balloon. Let's see pilking!*), did not help verb learning. Adults have also been shown to being able to learn verbs via syntactic bootstrapping (Fisher, 2002; Bunger, 2006), however, there is much less evidence. In Bunger's (2006) study, instructions were rather explicit. It is therefore an interesting question how implicit syntactic bootstrapping in adults could potentially be.

There are also other regularities in sentences beyond verb-argument structure which word learners can potentially benefit from. The syntactic frame (i.e., is the kinds of words which regularly occur before or after another word) has been found to also enable the learning of prepositions (Fisher, Klingler & Song, 2006). Function words such as determiners and pronouns can further provide notable assistance for word learning (Bernal, Lidz, Millotte & Christophe, 2007; Hoehle, Weissenborn, Kiefer, Schulz & Schmitz, 2004) because they are reliable signals for the grammatical category of the following word (a determiner is most often followed by a noun and a pronoun by a verb, Christophe et al., 2008). Phonological characteristics such as prosodic information can also provide cues about grammatical categories and therefore indirectly facilitate word learning (*phonological bootstrapping*, Morgan & Demuth, 1996; Gout, Christophe & Morgan, 2004; Christophe, Peperkamp, Pallier, Block & Mehler, 2004; Shatzman & McQueen, 2006). As Christophe et al., 2008 point out, prosody is potentially a very important cue because it can help to overcome the circularity of word and structure learning: Syntactic rules are easier learned if words are known (and it turns out that syntax learning does not happen before at least a number of nouns are learned) and words can be better understood if the sentence structure is clear (Gillette et al., 1999). Since phonological patterns can be understood independently of vocabulary and syntax, they are a potentially helpful cue. Millotte, Wales, Dupoux & Christophe (2006), for instance, presented evidence revealing that the prosodic phrase boundary, which is marked by variations in pitch and rhythm, helped French adults to disambiguate between two temporarily possible interpretations regarding the grammatical category of an unknown word in pseudo-French sentences. In all these experiments, instructions were either explicit or the experiment strongly suggested its aim.

2.2.3.2. The Role of the Linguistic Context in SLA Word Learning

While the bootstrapping studies cited above precisely reveal how local linguistic cues can be exploited, the sentential contexts used for testing are easier and more constrained than adults' language input tends to be. For decades, however, there has been a more general discussion about the importance and benefit of the linguistic context for implicit and explicit word learning in SLA, specifically about its role in making inferences during reading. However, although there is a relatively large number of studies on the influence of the linguistic context on explicit word learning (*lexical inferencing*; e.g., Nation, 1982; Carter, 1987; Prince, 1996; Barcroft, 2004), researchers do not entirely agree on its role. While some researchers argue that inferencing via the linguistic context is a crucial explicit word learning technique (Nation, 2001), many theorists have argued, that it is successful only when the learner has already achieved a certain level of vocabulary proficiency (*threshold hypothesis*, Schoonen, Hulstijn & Bossers, 1998) and when contexts are not too complex (Kaivanpanah & Alavi, 2008). Mondria and Wit-De Boer (1991) experimentally demonstrate that only contexts which strongly support a relatively constrained meaning for a novel word are helpful (*pregnant contexts*, e.g., *The gardener filled the X to water the plants.* vs. *I am looking for an X to finish my work.*). Schmitt (2000), comes to the conclusion that while making inferences can be a good strategy, the likelihood that it is successful is rather low. Teaching learners inferencing strategies, however, does enhance their performance. The relevance of intentionally using the linguistic context in spoken as opposed to written language comprehension has not been examined (Kaivanpanah & Alavi, 2008).

There are few attempts to examine the importance of the linguistic context in incidental (as opposed to explicit) learning settings, and those which exist, have often been criticized for a lack of real implicity (see Paribakht & Wesche, 1999 and Wesche & Paribakht, 2010 for discussions). Some researchers, however, found beneficial influences of the linguistic context in non-intentional reading (Read, 2000; e.g., Paribakht & Wesche, 1997; Paribakht & Wesche, 1999; Pulido, 2003). Paribakht & Wesche 1999, for instance, observed that university students used different ways of linguistic inferencing on sentence-level (word order, argument structure, parts of speech) and word level (morphological analysis) to find the meanings of novel words. While Paribakht & Wesche (1997) also found significant word learning after pure exposure to written text, they documented a reliably higher performance after reading combined with explicit vocabulary exercises. Most studies reveal, however, that, successfully inferring word meanings incidentally is, as in intentional inferencing, only possible when enough of the words are known, that is when learners are proficient enough for the level of the text (Schmitt, 2000; Paribakht & Wesche, 2010). Additionally, language novices only tend to make an effort to infer when a word is important for understanding: When they consider unknown words as unimportant, they often just ignore them (Read, 2000; Paribakht & Wesche, 1999). Surprisingly, there are almost no experimental examinations on the role of the linguistic context on incidental word learning based on spoken input comprehension but there are at least indications that the spoken linguistic context can be successfully exploited given that a word is salient, that is, important for understanding (Brown, 2008).

Learners' prior knowledge about states and relations in the world (*background knowledge*, *world knowledge*, or *context*) plays a special role regarding SLA word learning via the linguistic context. Given that this knowledge is mainly a result of experience, it has a higher beneficial effect for adults than children (MacWhinney, 2008). Research reveals that background knowledge boosts SLA vocabulary learning through reading, both incidentally (Paribakht & Wesche, 1997; Paribakht & Wesche, 1999; Pulido, 2003) and explicitly (e.g., Hamada, 2009; Kaivanpanah & Alavi, 2008). Interestingly however, there is a general lack, again, of studies on spoken language comprehension and also a lack of any controlled experimental studies about the precise influence of background knowledge on SLA word learning. These effects are, on the contrary, well studied in native language processing as will be discussed in Section 2.4.

2.2.4. Innate Constraints

People are known to be influenced by several innate constraints and adaptive biases which help word learning (at least in an early stage). For instance, children have a tendency to attend to whole objects instead of parts of objects (such as a rabbit rather than the rabbit's nose; *whole object bias*, MacNamara, 1972). Furthermore, while they are biased to understand adjectives modifying solid rigid objects as shape describing (*shape bias*, e.g., Landau, Smith & Jones, 1988), they attend more to color and texture when properties of animals are being talked about (Jones, Smith & Landau, 1991; Landau, 1998). Children also tend to think of basic-level categories (i.e., of an intermediate level of specificity) rather than more over- or under-specified classes ('dog' instead of 'poodle' or 'animal', Horton & Markman, 1980) and they prefer to extend object names to objects of the same kind rather than to something that is otherwise related (Markman & Hutchinson, 1984). Finally, children expect novel nouns to denote unnamed things (*Novel Name-Nameless Category Principle*, N3C, Golinkoff, Hirsh-Pasek, Bailey & Wenger, 1992) rather than to things which already have a name (*mutual exclusivity*, Markman & Wachtel, 1988; Hansen & Markman, 2009). The principle of mutual exclusivity is a critical constraint for infant word learning: When hearing a novel word, children are biased to assign this word to something which they cannot name yet because they prefer to build one-to-one referent-word mappings. It is not the case that the different constraints listed above are all at play all the time, rather, they have different weights and can be overridden in certain situations (e.g., Graham, Nilsen, Collins & Olineck, 2010).

Innate biases and constraints, again, have been predominantly studied within child language acquisition but still have been shown to have some relevance for adults: Adults use the principle of mutual exclusivity in a more flexible way (Yurovsky & Yu, 2008; Gangwani, Kachergis & Yu, 2010) but it does still influence their decisions to some extent (Au & Glusman, 1990; Golinkoff et al., 1992; Ichinco, Frank & Saxe, 2009). The shape bias is still present in adults (Landau et al., 1988; Landau, Smith & Jones, 1992; Markson, Diesendruck & Bloom, 2008) even if it's importance decreases and it looses out against other constraints such as attendance to function (Landau, 1998). One could imagine that these constraints may operate rather implicitly, also in adults, but this question remains unanswered.

2.2.5. Word Learning in SLA: Explicit vs. Implicit

While CSWL, social cues, SLCL, and innate constraints are necessarily applied incidentally in infants, this is less clear for adults. In general, learning words incidentally is not unproblematic in SLA because it is only successful when language novices receive enough exposures to the novel words (Elley, 1991; Ellis, 1997; Schmitt, 2000, Hulstijn, 1992). Spending time somewhere where the target language is actually spoken is thus very gainful (Milton & Meara, 1995), even for short periods of three to four weeks (Llanes & Muñoz, 2009). The most common alternative way to receive enough exposure is to read but also watching television can be helpful. Although implicit learning mechanisms are potentially very relevant for non-instructed situations, there are few studies that replicate this setting instead of focussing on very constrained and mainly text-based input. One compelling exception comes from Gullberg and colleagues (Gullberg, Roberts, Dimroth & Veroude, 2010; Gullberg et al., in press): They exposed Dutch adults to a minimal exposure of an entirely unknown foreign language (Chinese) in form of video-taped spoken weather reports (method based on Zwitserlood, Klein, Liang, Perdue, Kellerman & Wenk, 1994). Results reveal that, firstly, novel words could quickly be recognized and, secondly, word meanings could be learned if assisting gestures (pointing on the weather chart) were provided.

In contrast to infants, adults can moreover explicitly learn vocabulary in many different ways (Schmitt, 1997). Often language novices train themselves with the help of simple vocabulary lists including novel words paired with either translations or depictions (Read, 2000). On the one hand, it has been argued that rehearsing words this way is very effective and quick. On the other hand, theorists observe that techniques which involve more engagement of the learner usually lead to deeper processing (Schmitt, 2000; Sagarra & Alba, 2006). One example is Hulstijn's 1992 *Keyword Method*: Learners drew images for each novel word which combined the meaning of that word with the meaning of a similar-sounding word of the learner's mother tongue (e.g., Spanish 'table' *la mesa* with a picture of a 'messy' table). Learners also use a range of individual strategies to discover the meanings of words such as using dictionaries, making inferences from the linguistic context, and exploiting visual information (Schmitt, 2000). As discussed above, CSWL and SLCL can be also used explicitly.

The multi-modal context plays a potential role in both explicit and implicit vocabulary learning and potentially even more in non-instructed learning situations. Both visual cues and linguistic contexts have been shown to, firstly, influence adult learners incidentally and, secondly, to be exploited strategically (Schmitt, 1997). Importantly, the multi-modal environment does not only help to identify word meanings (incidentally and intentionally), there is evidence that learning with multi-modal cues enhances retention (in general: Clark & Paivio, 1991; Kelly et al., 1999; in vocabulary learning: Tight, 2010). Results of Tights' study reveal that noun learning and retention (tested with a multiple choice test and a translation task) is better when learners were exposed to cues in mixed modalities (visual, auditory, tactile) then when only one modality was used.

The role of both explicit and incidental word learning activities in SLA is an ongoing matter of discussion and investigation. Explicit learning leads to more rapid success but it is also time-consuming whereas implicit learning is slower but less effortful. While some extreme positions about the power of incidental word learning arose in the 1980's (Krashen's Input/Comprehension Hypothesis: Krashen & Terrell, 1983; 1985), nowadays most researchers agree that explicit vocabulary learning is inevitable, at least initially and to gain the core of a vocabulary (Cohen, 1987; Hulstijn, 1997; but see McQuillan & Krashen, 2008). Implicit word learning, on the contrary, has been considered to be a suitable way to learn less frequent words which are less important for a first basis (Schmitt, 2000). Schmitt (2000) emphasizes that explicit and incidental word learning are both necessary and are best seen as complementary.

While this discussion is important to understand how adults learn words, it is not the aim of this thesis to draw a conclusion about the weight of both kinds of learning. The mechanisms we are primarily interested in, namely CSWL and SLCL, can be either more incidental or more intentional in nature, as discussed above. As the study by Breitenstein, Kamping, Jansen, Schomacher & Knecht (2004) nicely demonstrates, CSWL can also well be adopted as an vocabulary trainer for healthy and aphasic adults. While Breitestein and colleagues found that learning in healthy adults was better when feedback was provided, this was not necessary for learning success. The authors note, however, that some degree of awareness was also involved in the no-feedback condition because the learning method was transparent. In our experiments, learning is somewhere between incidental and explicit word learning: While the instructions are only partly explicit, learners tend to quickly understand that the subject of the experiments is to learn novel words. The way learners are trying to find the meanings of the novel words is not at all suggested to them and oral self-reports reflect only partial awareness of the mechanisms at play. We therefore think that conclusions for a theory of implicit SLA (in non-instructed contexts) can be drawn from our results. Additionally, however, our paradigm allows us to make implications for explicit vocabulary training.

2.2.6. Weighting and Interaction of Learning Mechanisms

As shown above, there is a variety of potentially available information cues and learning mechanisms which naturally help word learners. However, it seems very unlikely that word learners, both infants and adults, rely on only one of these mechanisms. That means that to investigate the real contribution of diverse word-learning mechanisms, it is necessary to study them in a combined fashion. Assuming that mechanisms work separately and isolated may one the one hand over-complicates the situation because it under-estimates people's solution-solving capacities (i.e., using several cues and mechanisms in parallel to find a solution). On the other hand, different mechanisms potentially influence each other in complex ways and may, for instance, decrease each other's power. However, surprisingly little research has addressed the weighting and interaction of diverse word-learning mechanisms, either in children, or in adults.

While some researchers attribute early word learning to only (or mainly) one of the groups of mechanisms presented below (e.g., social-pragmatists to social cues) others claim that different ways of word learning interact (Hollich, Jusczyk & Brent, 2000). Moreover, the weight of different cues seems to underlie developmental changes (Baldwin, 1993; Brandone, Pence, Golinkoff, Michnick & Hirsh-Pasek, 2007; Shepherd, 2010). Hollich et al. (2000) point out, for instance, that while associative and perceptual (attentional) cues are more influential in young infants, social cues take priority over them at some point (at about 24 months of age). This discussion is important and interesting, however, it still does not consider the full range of potential influences.

Some important findings come from a study presented in Gillette et al. (1999). They found that combined syntactic information (verb frame), lexical information (nouns) and scene information (video) can result in better verb learning in adults than only one of these cues (see also Snedeker & Gleitman, 2004). While these results demonstrate compellingly that adult learners can combine different multi-modal cues which complement each other, they do not shed light onto their interplay and weighting because all information is clear and cues are in accordance with each other. It is unclear, for instance, how learners would react if distinct cues are ambiguous or even contradictory. Crucially, a study replication with infants did not reveal the same results and some conditions with less information provided were better than others with more information (Lidz, Bunger, Leddon, Baier & Waxman, 2010). The authors attribute this finding to infant's limited processing capacities and conclude that there is a tradeoff between the informativeness of a cue and the processing resources it requires. That means, importantly, that adult learners may be advantaged in that they are better at integrating different cues into their comprehension and learning.

Some researchers have investigated the interplay of CSWL and some cognitive constraints, in particular, the principles of contrast (Kachergis, Yu & Shiffrin, 2010a) and mutual exclusivity (Ichinco et al., 2009; Yurovsky & Yu, 2008; Gangwani et al., 2010). Results of two CSWL experiments presented in Gangwani et al. (2010), for instance, reveal that adults were able to at the same time learn two labels for one object: one concrete object name and a category name. These findings are insightful because they suggest a more moderate version of mutual exclusivity. However, many questions about how the mechanism of CSWL interacts with other cues remain open. We will address this important lack of empirical research and discussion in our experiments.

2.2.7. Interim Summary

To summarize Sections 2.2.1 to 2.2.6, while there is convincing evidence that infants learn words with the help of observational learning, analysis of co-occurrences (CSWL), social cues, innate constraints, and the linguistic context (SLCL), only some of these mechanisms have been convincingly shown to operate in SLA word learning, some more incidentally, some more explicitly: Adults perform very well at CSWL and they have been shown to successfully use SLCL. Using prior knowledge about linguistic structures and the world, which is helpful for SLCL, is a source that adults can benefit from better than infants because they have naturally collected more experiences in their lives (MacWhinney, 2008). Some of the mechanisms which have been demonstrated to be important for infants (e.g., learning via social cues such as gesturing) may be under-estimated in their role for sword learning in adults because second language acquisition has rarely been studied in non-instructed and multi-modal contexts. Adults can additionally apply many different explicit word learning strategies, which are as important as implicit learning mechanisms for SLA. There are additionally some indications that adults can better benefit from multiple cues in parallel than children (see Lidz et al., 2010), however, not much research on the interplay of different cues has been done. Taken together, the review above reveals that research, in particular on the interaction of word-learning mechanisms, leaves many questions unanswered, some of which will be addressed in this thesis: Firstly, we will study CSWL in adults in more natural contexts, secondly, we will investigate SLCL in adults with spoken sentences and examine the role that world knowledge plays in a controlled manner, and, thirdly, we will examine the interaction of CSWL and SLCL.

2.3. Sentence-Processing Mechanisms

To study situational influences on word learning - with words embedded in sentences and sentences embedded in visual contexts - it is also necessary to consider how sentences are processed and to identify which characteristics of sentence processing may be relevant for word learning. The fact that spoken sentence comprehension is incremental, predictive, multi-modal, and constraint-based potentially plays a particular role. We will now elaborate on these characteristics and discuss their scope and specific role in L2 sentence processing.

2.3.1. The Nature of L1 Sentence Processing

To understand a spoken sentence, people have to generate a mental representation of what is being described, that is, they have to build a mental model (or discourse model; e.g., Johnson-Laird, 1983; Garnham, 1981; Marslen-Wilson & Tyler, 1987). While this sounds intuitive and plausible, complex architectures and mechanisms are involved in this process and there are many proposals regarding how it can be best described and modeled (see Gernsbacher & Kaschak, 2002 for an overview). Two characteristics, however, are largely agreed on: Firstly, language comprehension is rapid, highly *incremental*, and *predictive* (Marslen-Wilson, 1973; Pickering, Clifton & Crocker, 1999; Van Gompel & Pickering, 2007) and, secondly, language comprehension is *multi-modal* (Tanenhaus & Trueswell, 1995; Knoeferle, Crocker, Scheepers & Pickering, 2005; see Crocker, Knoeferle & Mayberry, 2010 for a review).

Incrementality with regard to sentence processing means that people interpret linguistic input word by word, making sense of the pieces of the sentence as soon as they are encountered, rather than waiting until the sentence is finished. There is convincing evidence supporting this claim, for instance from studies investigating the interpretation of temporarily ambiguous sentences (e.g., Rayner, Carlson & Frazier, 1983; Ferreira, 1986; Trueswell, Tanenhaus & Garnsey, 1994): In a sentence such as (1), people tend to encounter difficulty (they are *garden-pathed*, Bever, 1970) when processing the verb *sneezed* because they interpreted the verb *raced* as the past tense verb of the matrix clause rather than the past participle verb of a reduced relative clause.

(1) The camel raced past the fountain stopped suddenly.

Sentence processing is not only incremental but also rapid enough to enable adults to predict upcoming parts of the sentence (Tulving & Gold, 1963; Van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005; Van Berkum, Brown & Hagoort, 1999; Federmeier, 2007). Staub and Clifton (2006), for instance, provide evidence for linguistic prediction on a syntactic level: The conjunction *or* is read faster when the preceding context contained *either* (2) than when it did not (3).

(2) The team took either the train or the subway to get to the game.
(3) The team took the train or the subway to get to the game.

It is interesting to think about the effect that prediction in language processing has. It has been shown that it makes understanding more rapid and effective (e.g., Schwanenflugel & Shoben, 1985; Staub & Clifton, 2006; see also Kveraga, Ghuman & Bar, 2007). Moreover, predicting has been argued to be helpful *to compensate for problems with noisy or ambiguous input* (Pickering & Garrod, 2007: 105). It is therefore possible that predicting also has a special role when language input is partially unknown, that is, when people encounter unknown words: Given that a language novice has an expectation about the semantic content of an upcoming word but does not know the phonetic form of that word (because he has not yet learned it), he may be advantaged in quickly mapping this semantic content to that novel word as soon as it is available.

Additionally, sentence processing is multi-modal: The interpretation a person recovers is dependent not only on linguistic information on various levels such as syntax, semantics, pragmatics, discourse, and prosody but also on non-linguistic cues such as the visual context and world knowledge (Altmann & Kamide, 1998; Sedivy, Tanenhaus, Chambers & Carlson, 1999). Integrating multi-modal information sources incrementally plays a potentially essential role for SLA word learning because it means that adults can use various cues immediately when encountering a novel word.

Convincing evidence for both incrementality and prediction on the one hand and multi-modal source integration on the other hand comes from studies using the visual world paradigm. In this experimental procedure, participants are exposed to visual scenes and spoken sentences while their eye-movements are recorded (Tanenhaus & Trueswell, 1995). One influential study, by Altmann & Kamide (1999), reveals that when adults hear a *restrictive* verb (a verb which constrains the semantic category of its direct object), they immediately prefer to look at those objects in a scene which match the verb's semantic category. When hearing Sentence (4) while seeing the scene in Figure 2.3, for instance, people tended to look at the cake after the offset of the verb (*eat*), that means, even before it was named (*anticipatory eye movements*). In contrast, when hearing Sentence (5), all objects were looked at equally often during the verb (*take*). The authors concluded that participants used their world knowledge, the visual scene, and the sentential information (the restrictive verb) to rapidly infer that something edible will be named next and exploited the visual world to identify the only edible object, the cake. During the postverbal NP (*the cake*), participants preferred to look at the cake regardless of the verb (i.e., in both Sentences 4 and 5; *referential eye movements*).

(4) The boy will eat the cake.

(5) The boy will move the cake.

There are also other cues that adults integrate into their language understanding. Dynamic event knowledge and linguistic tense information, for instance have also been shown to rapidly influence L1 sentence comprehension (Knoeferle & Crocker, 2007; Altmann & Kamide, 2007). Altmann and Kamide (2007), for instance, found evidence that participants' eye movements were simultaneously influenced by verbal restrictions, their world knowledge and event knowledge: The sentence fragment *The cat has killed* elicited more looks to a

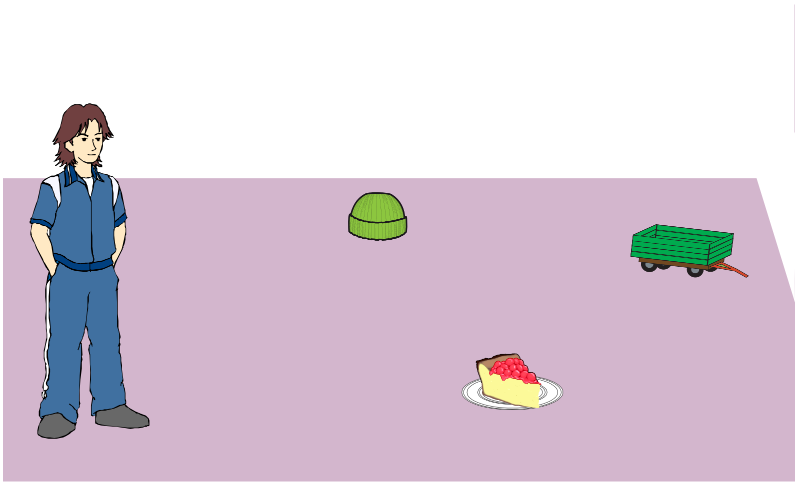


Figure 2.3.: Example scene visual-world paradigm

pile of feathers than to a group of mice which were still alive (see also Chambers, 2002, Altmann & Kamide 2004, and Altmann & Kamide, 2009 for further evidence).

While this hypothesis has not yet been experimentally examined, all these multi-modal cues may influence word learners in a similar and similarly rapid way. In a setting such as in Altmann and Kamide (1999), for instance, a learner would hear a transitive sentence and concurrently see a scene which contains, among other things, the referents which the nouns in the sentence denote. Given that the learner understands the verb but not the direct object (e.g., *The boy will eat the dax.*), for example, she could use the verbal constraints, her world knowledge, and the visual scene to rapidly identify the meaning of the novel noun (i.e., *cake*), potentially even predictively.

2.3.2. L2 Sentence Processing

It is particularly important for this thesis to examine how non-native speakers process language. Firstly, it needs to be examined whether there are particular characteristics in L2 sentence processing which might effect participants in our experiments. Secondly, it is crucial to investigate whether L2 sentence processing is similar to L1 sentence processing in its incrementality, prediction, and multi-modality. Finally, it is interesting to specifically study sentence processing when learners encounter (lexical) knowledge gaps.

2.3.2.1. Characteristics of L2 Sentence Processing

While there is considerable agreement that L2 learners have native-like abilities regarding argument structure (thematic roles) and lexical-semantic-pragmatic information during parsing (e.g., Frenck-Mestre & Pynte, 1997; Williams, 2006; Lew-Williams & Fernald, 2007), learners' abilities regarding syntactic analysis have been a matter of debate. Some researchers suggest that adult second language learners are less strongly guided by syntactic information than native speakers and rather rely on lexical-semantic cues (Clahsen & Felser, 2006; Felser & Clahsen, 2009). Papadopoulou & Clahsen (2003), for instance, present supporting results which reveal that while relative clause attachment in native speakers of Greek is biased by structural and lexical factors, processing in L2 (with L1 German, French, or Russian) is exclusively determined by lexical constraints. Clahsen and Felser (2006) come to the conclusion that the syntactic representations that language novices generate during processing are less detailed and shallower than in native speakers (*shallow structure hypothesis*): When parsing a sentence, L2 learners generate predicate-argument structure based on thematic roles and other lexical-semantic information whereas they (partly) ignore structural relations such as syntactic hierarchies. However, limitations in syntactic processing have only been found for certain phenomena (long-distance dependencies, e.g., in wh-questions; Marinis, Roberts, Felser & Clahsen, 2003; Felser & Roberts, 2007) while other structures have been shown to be comprehended native-like (local syntactic phenomena such as subject-verb agreement and gender concord, Ojima, Nakata & Kagigi, 2005; Sabourin & Haverkort, 2003).

Clahsen & Felser (2006) argue that sentence comprehension is slower in L2 than L1. They assume that this is possibly due to cognitive resource limitations or a lack of automaticity. While this suggestion may be supported by the data resulting from some ERP-studies, reflecting an absence of certain components (namely, an early anterior negativity, LAN; e.g., Mueller, 2005), this issue is not yet clarified in a satisfying way (also because researchers do not agree on the role of the LAN; Clahsen & Felser, 2006). In general, many similarities between L1 and L2 sentence processing have been found in ERP studies (Mueller, 2005; Mueller, 2009). Kotz (2009), summarizing ERP and fMRI research, argues that L2 proficiency plays a significant role for brain activities: Many studies reveal that with increasing proficiency, L2 brain responses become increasingly similar to L1.

While L2 sentence processing may be slower than L1 sentence processing, there is convincing evidence that it is incremental and predictive as well as sensitive to semantic-pragmatic and multi-modal information (SLA: Boland, Tanenhaus, Garnsey & Carlson, 1995; Fender, 2001; Williams, 2006; artificial language: Wonnacott, Newport & Tanenhaus, 2007; Amato & MacDonald, 2010). Wonnacott showed that adult learners of an artificial language process sentences incrementally and even predictively after three days of training (possibly even earlier): Eye-movements revealed that learners anticipated referents based on verb-specific biases. More than that, Amato and MacDonald (2010), also using an artificial (though typologically implausible) language, reported that learners were able to combine semantic constraints on the sentence-final object which result from the combination of subject and verb. These results are very similar to the findings for native speakers obtained by Kamide, Altmann & Haywood (2003). Juffs (2009) presented data revealing that both native and advanced non-native speakers of English run into the same garden path effect when encountering temporarily ambiguous sentences such as (1): Readers integrate the postverbal object (*the water*) into the verbal phrase (*drank*) of the subordinate clause (*After Bill drank*) because to drink is frequently used as transitive verb (see also Roberts & Felser, 2011).

(1) After Bill drank the water proved to be poisoned.

Boland, Tanenhaus, Garnsey & Carlson (1995) showed, importantly, that this effect interacts with the plausibility of an NP as a verb's direct object in both native and non-native speakers. This result nicely reveals how foreign language learners consider both structural and semantic-pragmatic information (world knowledge) into their comprehension, just as native speakers do.

Taken together, given an appropriate relation between input complexity and learner's proficiency, adult SLA sentence processing resembles adult native language processing to a large extent in that it is generally quick, incremental, predictive, constraint-based, and multimodal. The processing of complex syntactic structures such as long-distance dependencies may be different in L2 processing than L1 processing but this limitation will not be relevant for out own experimental investigations. Finally, it is interesting to note that L2 sentence processing may even be influenced more strongly by syntactic-semantic constraints (argument structure) and lexical-semantic cues than native comprehension.

2.3.2.2. Compensation Strategies in L2 Sentence Processing

Language processing, in particular listening, in beginning L2 learners is problematic: Not only are language novices faced with partly unknown words and syntactic structures, they additionally have to cope with the speed and uncontrollability of the input and their limitations in working memory (Field, 2004; Jones & Plass, 2002; Vandergrift, 2004; Chang & Read, 2007). It is therefore all the more important to consider information from the multi-modal context because that may compensate for knowledge gaps on-line, implicitly and explicitly (Rubin, 1994 and references therein; Chang & Read, 2007; Vandergrift, 2004; Vandergrift, 2007). This process potentially, but not necessarily, contributes to word learning. Vandergrift (2007) argues that the visual context, the linguistic context (including the discourse), and background information (world knowledge) play a particularly important role when learners encounter knowledge gaps:

> 'Compensatory mechanisms - contextual, visual or paralinguistic information, world knowledge, cultural information and common sense - are used strategically by L2 listeners to compensate for their inadequate knowledge of the target language.' (Vandergrift, 2007: 193)

Visual support has been shown to be a useful cue in different environments (Meskill, 1996; Mueller, 1980; Seo, 2002; Ginther, 2002; Jones & Plass, 2002). Ginther (2002) presented evidence for the positive effect of pictures on comprehension when they contained information complementary to the spoken input. Results from Jones & Plass (2002) reveal that multimedia environments can be gainful for L2 listeners: Joined pictorial support and written annotations led to better recall and learning of vocabulary (see also Jones, 2006 and Chang & Read, 2007 for similar studies and results). Comprehended information (and vocabulary retention) was moreover longer-lasting when pictures had been presented than when no pictures had been presented. There is also some evidence for the positive influence of gestures on SLA sentence comprehension (Sueyoshi & Hardison, 2005). Finally, some researchers examined the role of dynamic visual elements such as in multi-media sources (e.g., movies) for L2 comprehension (Meskill, 1996; Secules, Herron & Tomasello, 1992). Secules et al. (1992) presented a study comparing two groups of American students of French which attended different classes, one with and one without video sessions once a week. The authors found that students from the video group performed better in comprehension tasks (video comprehension) at the end of the semester. However, studies investigating the effect of video material for language learning tend to be relatively uncontrolled and hence only provide rather general support. In general, while there is some evidence for the positive effect of visual context on SLA spoken sentence comprehension, the number of well-controlled studies and is surprisingly low and on-line measurements are not commonly used.

When processing SLA spoken input, prior knowledge and experiences are automatically activated, as in L1 comprehension. This schema activation is a crucial process because it is a further way to bridge comprehension gaps via inferencing (Meskill, 1996). Making inferences and predictions is as essential in L2 listening as it is for vocabulary learning (Rost, 2002; Hulstijn, 2003; Hinkel, 2006). Most studies, however, examine reading rather than listening. There is some empirical support for the very plausible assumption that background knowledge also helps to bridge knowledge gaps in L2 listening (Long, 1990; Chiang & Dunkel, 1992) but carefully controlled (and on-line) studies about the integration of this kind knowledge as a means of compensation are missing.

Finally, it is interesting to note that compensation strategies have been demonstrated to be teachable (Vandergrift, 2004; Field 2007; Field, 2007; Hulstijn, 2003; Goh, 1997; Mendelsohn, 2001). Hinkel 2006 presented data which reveal that equipping learners with prior knowledge before listening (*activation schema*) prepares them to make predictions and inferences while listening enhance listening comprehension. This, again, suggests a practical use of investigating language learners' use of compensation and motivates on-line studies of L2 spoken sentence processing.

2.4. Constraint-Based Language Learning and Processing

The review in Chapter 2 suggests that there are some interesting parallels between word learning and language processing (both in L1 and L2): Both word learners and sentence comprehenders are guided by different and interacting information sources which are based on various linguistic levels, the visual context, and people's knowledge and experience. This multi-modal and interactive nature of processing is well captured by constrained-based accounts. For sentence processing, there are some well known constraint-based models (e.g., MacDonald, Pearlmutter & Seidenberg, 1994; McRae, 1998). While it is a matter of debate whether syntactic information is given (temporal) priority for interpretation (e.g., Frazier, 1979; Fodor, 1983), there is convincing evidence that cues do interact, at least to some extent (e.g., visual world studies cited in Section 2.3.1; Trueswell et al., 1994; Spivey, Tanenhaus, Eberhard & Sedivy, 2002; Chambers, Tanenhaus & Magnuson, 2004; Knoeferle et al., 2005; McRae & Matsuki, 2009). Given that the precise role of syntax is not relevant to this thesis, this issue will not be further discussed.

There are also attempts to describe both processing and learning of L1 and L2 with one theory (Harrington, 2004; MacWhinney, 2005; MacWhinney, 2008). The *Unified Model* by MacWhinney (2005; 2008) is a modification of his constrained-based, interactive Competition Model (Bates & MacWhinney, 1982; MacWhinney, 1987). The main characteristic of this unified account is that it is constraint-based, interactive, parallel, and distributed: Information from different sources is integrated simultaneously into interpretation, including lexical-semantic cues (e.g., plausibility, animacy), syntactic cues (e.g., word order), morphological cues (e.g., case markings), morphosyntactic cues (e.g., subject-verb agreement), and prosodic cues (e.g., lexical stress). These cues or constraints, which have certain weights or strengths subject to permanent updating, can converge or compete; when in competition, each cues's influence is determined by its current *cue strength* (equals the *cue validity* in the Unified Model). While the cue validity is considered as a product of *cue availability* and *cue reliability* in children, it is a direct function of cue reliability for adults (MacWhinney, 2005; 2008). When processing the ungrammatical English sentence *The eraser push the dogs*, for instance, people's interpretation about which NP is the subject of the clause is influenced by three different cues: sentence position, subject-verb agreement, and animacy (plausibility). Subject-verb agreement and plausibility support interpreting the NP2 *the dogs* as subject, however, sentence position suggests that the NP *the eraser* is the subject.

Given that sentence position is the strongest cue because it is very reliable, even when in conflict with other cues, most English speakers prefer to choose the second analysis.

While MacWhinney does not explicitly illustrate how his model captures the encountering of unknown words, it is relatively straight forward to just apply his general idea of cue competition to this scenario. That means that when language novices encounter unknown words, all available cues are activated in order to find a meaning for the word. For instance, when an adult learner hears the novel noun *dax* embedded in the sentence *The dax is handing the cake to the girl* and at the same time he sees a dog in front of a cake, he has to weight the conflicting cues of visual context (supporting the hypothesis that *dax* means dog) and plausibility (supporting the hypothesis that *dax* refers to a human). Both cues could also complement each other, for instance when the visual scene depicts a clown and a dog in front of the cake rather than only the dog. In that case, both constraints, plausibility and visual information jointly support the hypothesis that *dax* denotes the clown. As discussed above, there are only few studies investigating the interaction of different learning cues and mechanisms and therefore there is little evidence that word learning is interactive. However, studies such as Gillette et al. (1999) and those by Yu and colleagues, discussed in Chapter 2.2.6, support an interactive account.

2.5. Summary and Motivation

Word learning is a crucial but demanding task. This chapter has reviewed evidence that people overcome its difficulty through their ability to use multi-modal information sources in various ways. In particular, children and adults gain the command of a language's vocabulary via observing the visual context and using cross-situational co-occurrence frequencies (CSWL), considering social cues, making inferences based on both the linguistic context and people's knowledge about linguistic regularities and the world (SLCL), being guided by innate constraints, as well as via using explicit strategies of word-meaning identification and memorization. While research provides evidence for the general validity of all these mechanisms, some may be more important for child word acquisition (specifically, innate constraints and social cues), whereas others are certainly only relevant to SLA (explicit strategies). Two extremely powerful word-learning mechanisms, CSWL and SLCL, have been shown to be important for infants and adults. Notably, however, research on these word-learning mechanisms is usually based on unnaturally isolated experimental stimuli and tasks. A few studies suggest that different mechanisms also interact, however, this issue is highly under-researched. As demonstrated, the way adults process both L1 and L2 sentences, that is, incrementally, predictively, and multi-modally contributes significantly to solving the word learning puzzle: Learners can rapidly use available information from all different sources. The differences between L2 sentence processing and L1 sentence processing

(L2 processing is potentially a bit slower and may in some cases lead to shallower syntactic analyses), are not relevant to word learning. Except its general importance for sentence comprehension, the multi-modal context plays a particular role when L2 comprehension is disturbed because it can compensate for knowledge gaps. Finally, we argued that the nature of both word learning and sentence processing can be accounted for by interactive constrained-based models, specifically, by the Unified Model by MacWhinney (2005; 2008).

In the remainder of this thesis, nature and interplay of two of the described word-learning mechanisms are investigated, namely CSWL and SLCL. Specifically, the kind of SLCL we study in our experiments is motivated by the sentence processing study by Altmann and Kamide (1999), presented above. The rationale is that when learners hear a novel noun following a restrictive verb such as *eat*, they use their world knowledge to infer that the noun is likely to refer to something edible which motivates them to identify edible objects on a visual scene. There are several reasons to select CSWL and SLCL: Of the group of mechanisms potentially working in non-instructed environments (rather implicitly), they are those two which are most clearly relevant for SLA, as identified in the literature review. Moreover, the interplay of both is potentially very interesting because, while resembling one another in being visually situated, they are very different in nature. CSWL is an associative and bottom-up process: Concurrent linguistic and visual stimuli are mentally linked to each other and the strength of these links is then increased or decreased depending on further evidence or counter-evidence across episodes. That means that no prior knowledge about linguistic or non-linguistic regularities is per se needed to conduct CSWL, all that is necessary is the multi-modal (linguistic and visual) context and a cognitive system memorizing connections between the two. CSWL has been further argued to be a parallel, incremental, and probabilistic means of word learning (Vouloumanos, 2008; Yurovsky et al., 2010). SLCL, on the contrary, works more top-down: Language novices need to integrate their knowledge about both linguistic structures on different levels and about the world. SLCL is potentially also more deterministic than CSWL because referents can be clearly identified within one situation. The following expectations about the interaction of CSWL and SLCL drive the design of our experiments: Firstly, given that both mechanisms provide information from different perspectives (top-down vs. bottom-up), it should be possible to use them in a complementary and joint way to identify word referents. Secondly, given that they are both powerful, they should reduce each others' influence when they are in conflict: On the one hand, SLCL may be more dominant than CSWL because it is more reliable for language novices: SLCL is based on people's prior knowledge, while CSWL is not. On the other hand, CSWL may effect word learners stronger than SLCL on an associative (possibly unconscious) level. Thirdly, given that SLCL is more deterministic in nature than CSWL, it might also be expected to suppress CSWL (in certain situations).

We will now present five psycholinguistic experiments, all of which are based on teaching participants a miniature semi-natural language in a stepwise procedure. In particular, there were three main phases in our experiments: Participants were familiarized with a set of verbs, then they learned novel nouns, and, finally, noun knowledge was assessed. Crucially, we used a novel design for noun learning: Nouns were embedded in linguistic contexts and these linguistic contexts were situated in visual environments. This constitutes a naturalization of the setting used in learning experiments and enabled us to study interactions of learning mechanisms based on different modalities. In Chapter 3 (Experiments 1 to 3), CSWL and SLCL are investigated as complementary word-learning mechanisms. In particular, Experiment 1 was conducted firstly, to investigate CSWL in naturalized situations (with nouns embedded in SVO sentences and referents as parts of scenes), secondly, to study the influence of SLCL on on-line sentence comprehension and noun learning and, thirdly, to investigate CSWL and SLCL as complementarily applicable. We hypothesized that participants use both mechanisms in a complementary way to learn the set of nouns and that SLCL boosts the on-line comprehension and learning of (syntactic) object nouns. The main aim of Experiment 2 was to study the effect of verb-based anticipation of referents on noun learning. Our main hypothesis was that nouns would be learned better if it is possible to anticipate the referent via a restrictive verb (because the verb preceded the noun) than if it is not possible (because the verb followed the noun). In Experiment 3, we re-addressed the question whether SLCL boosts noun learning. We also examined the interaction of CSWL and SLCL when both are applicable in a complementary way with regard to identifying the meaning of one word. We hypothesized that noun learning would be easier when learners can apply both CSWL and SLCL than when only CSWL is available and that learning via pure SLCL would be better than learning via pure CSWL. Chapter 4 (Experiment 4) addresses the interaction of both cues when they are in conflict, that is, when SLCL supports another meaning than SLCL. We hypothesized that both mechanisms would influence learners' decisions about equally strongly. The aim of Chapter 5 (Experiment 5), finally, is to examine the interplay of CSWL and SLCL as independently applicable, to define the underlying processing differences of both mechanisms, and to evaluate the retention of the learned nouns (as assessed by a vocabulary test one day after learning). We hypothesized that while CSWL works probabilistic and parallel, SLCL is more deterministic and category-based and that SLCL blocks CSWL when both mechanisms are independently applicable. Moreover, we hypothesized that learning via both mechanisms would lead into long-term learning.

3. CSWL and SLCL as Complementary Learning Mechanisms

In Chapter 2, two mechanisms of visually situated word-learning mechanisms were identified: CSWL and SLCL. In this chapter, we report the first three of a series of five experiments exploiting a naturalized learning setting. The first aim of these three experiments, presented in this chapter, was to determine whether CSWL is utilized in such a complex setting (i.e., with words which are embedded in a linguistic context and visual referents which are parts of scenes). This is an important question given that CSWL has so far only been studied within highly simplified learning environments. Our second aim was to examine whether SLCL helps language novices to identify (Experiments 1 to 3) and learn (Experiment 3) novel nouns. Specifically, we examined the influence of verbal constraints together with the visual context and learners' world knowledge on noun identification and learning: Motivated by the combined evidence for verb learning via argument structure (syntactic bootstrapping, e.g. Bunger, 2006) and anticipation of referents elicited by restrictive verbs in on-line sentence comprehension studies (Altmann and Kamide, 1999), we expected that noun learners would benefit from these cues. We further investigated whether CSWL and SLCL interact in a complementary way, both with regard to learning a set of novel nouns (Experiments 1 to 3) and for learning a single noun (Experiment 3). As explained in Chapter 2, we expect that CSWL and SLCL can jointly assist word learning, hypothesizing that, firstly, adults are generally capable of using multiple mechanisms in parallel and, secondly, CSWL and SLCL utilize different cognitive resources (bottom-up versus top-down). Finally, we addressed the role of verb-based anticipation of referents on object-noun learning (Experiment 2), hypothesizing that such predictions may improve word learning (as discussed in Chapter 2).

3.1. Experiment 1

3.1.1. Motivation

Experiment 1 was set out to evaluate CSWL in a more situated and natural setting than has been common practice: With nouns embedded as parts of sentences and referents embedded in scenes. While this scenario may complicate noun learning because it adds complexity, it may also simplify the task given that the linguistic context and the visual scene potentially

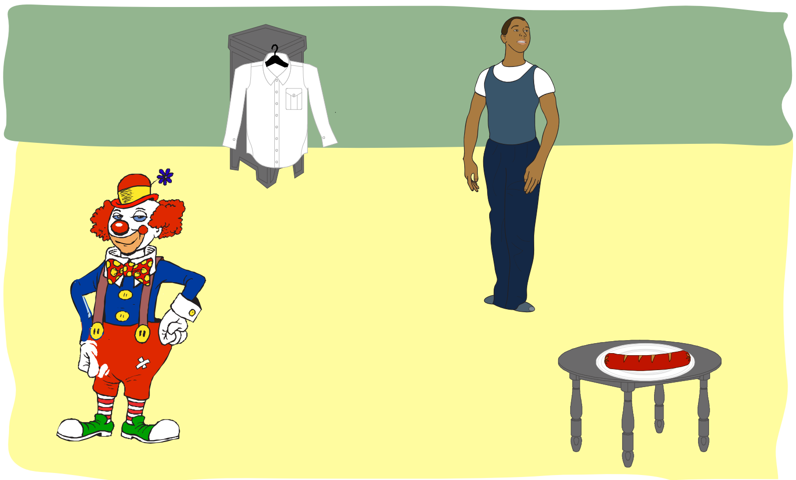


Figure 3.1.: Example item Exp. 1
Si badut bermamema si worel.
'The clown will eat the sausage.'

provide useful cues. Studying noun learning in this way additionally enabled us to examine the influence of SLCL. In particular, we investigated the impact of restrictive verbs (together with both the visual context and learners' prior knowledge about language structures and the world) on the comprehension and learning of nouns in post-verbal sentence position. Our hypothesis was that this kind of SLCL would be a powerful mechanism for both identifying referents on-line and learning nouns. We moreover aimed at studying the interplay of CSWL and SLCL when both are complementary, hypothesizing that both mechanisms can be jointly used by adults. A final aim of Experiment 1 was to investigate sentence processing in adult language learners on-line. Specifically, we hypothesized that learners would process sentences in a similar way as native speakers, that is, incrementally, predictively, and multi-modally.

The learning procedure in our experiments was stepwise: First, participants were familiarized with a set of verbs. Second, they were exposed to noun-learning trials, and, third, they conducted a forced-choice vocabulary test and a sentence judgement test. Eye-movements were recorded during noun learning to examine learners' on-line referent identification. For each noun-learning trial, one spoken SVO sentence was paired with a visual scene (see example in Figure 3.1). Nouns were syntactic subjects (example: *si badut*) or objects (example: *si worel*) in these sentences. Each subject noun referred to one of two agent characters and each object noun referred to one of two inanimate objects depicted in the scene. While the verb's selectional restrictions could be used to identify the more likely referent of the object noun (example: the sausage), subject nouns were unconstrained by the verbs (beyond the general bias towards animates). That means that object nouns could potentially be identified immediately, using SLCL, whereas subject nouns had to be learned cross-situationally. Assuming that participants would conduct both CSWL and SLCL, we hypothesized, firstly, that on-line identification of object-noun referents would be easier for learners than on-line identification of subject-noun referents and, secondly, that learning of object nouns would be more successful than learning of character nouns. Furthermore, we hypothesized that L1-like anticipatory eye-movements following a restrictive verb (Altmann and Kamide, 1999) would occur.

3.1.2. Methods

3.1.2.1. Participants

32 German native speakers took part in Experiment 1 (for a reimbursement of €5), eight of which had to be excluded due to technical problems and experimenter error. The data of 24 participants was analyzed (17 female, 7 male, aged between 22 and 47, average age 30, mainly students from various disciplines).

3.1.2.2. Materials and Procedure

The overall task of the experiment was to teach participants a miniature semi-natural language which was based on Indonesian. In the experimental instructions, this language was introduced as a language spoken by Indonesian circus people. It consisted of six restrictive verbs, twelve nouns, and one article that was used as the determiner for all nouns (*si*). All 12 twelve nouns were two syllables long, six nouns referred to animate characters ('clown', 'acrobat', 'ballerina', 'magician', 'circus director', 'boxer'), three to food items ('sausage', 'corn', 'mushroom'), and three to clothing items ('shirt', 'trousers', 't-shirt'). Verbs all consisted of two syllables plus a suffix (*-mema*) and belonged to one of two meaning categories (food: 'eating', 'barbecuing', 'salting'; and clothing: 'drying', 'ironing', and 'sewing'). Word order was subject-verb-object (SVO). To avoid any effects due to particular associations, two different vocabulary-meaning mappings were created. Half of the participants saw the one mapping (List 1), half the other (List 2).

The experiment comprised three main phases: isolated verb learning, eye-tracker preparation, and verb repetition (Phase 1), sentence comprehension and noun learning (Phase 2), and vocabulary testing (Phase 3). All stimuli were presented on a computer screen and via computer loudspeakers, using Experiment Builder software. The entire experiment lasted approximately 30 minutes.

Verb Learning In Phase 1 (verb learning), participants were familiarized with the six verbs (see 3.2). Static pictures showing actions were presented on the computer screen,

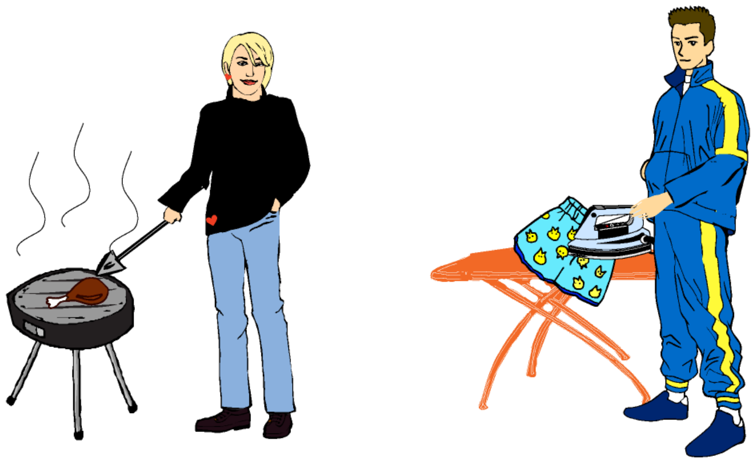


Figure 3.2.: Verb learning Exp. 1
left: *mankemema* ('barbecue')
right: *melimema* ('iron')

together with spoken verbs played back over loudspeakers. Instructions explicitly included, firstly, the information that the spoken words are names for the visible actions and, secondly, the request to memorize mappings as well as possible. Verb learning was not aimed to simulate a realistic learning situation but it was simply meant to make participants familiar with these verbs. Two pictures were used for each verb, differing in both the actors and the manipulated objects. Each of these two combinations was repeated five times (i.e., 10 presentations per verb). The order of presentation was randomized in the middle part but controlled in beginning and end because observations from pilot data suggested this structure to ensure better learning performance. In particular, in the controlled parts, verbs were repeated in a fixed order (Verb 1, Verb 2, Verb 3, Verb 4, Verb 5, Verb 6, Verb 1-6, Verb 1-6 etc.).

Next, verbs were tested: Participants were presented a picture and asked to pronounce the matching verb (pictures were object-character combinations that had not been used for the learning phase). Each verb had to be named two times, feedback was provided. Before the second learning phase, the head-mounted eye-tracker (EyeLink II) was prepared: First, it was adjusted to the participant's head size and one camera was positioned such that the participant's dominant eye could be recorded. Next, the eye-tracker was calibrated. After calibration, verbs were quickly repeated: Each of the 12 test pictures was shown again and named by the experimenter in case the verb could not be produced by the learner.

Noun Learning In Phase 2 of the experiment (noun learning in sentence contexts), participants were simultaneously exposed to static scenes and spoken sentences. Each scene depicted simple indoor or outdoor scenes with two characters and two objects (one clothing item and one food item). See Figure 3.1 for an example. While participants were already familiar with verbs, they were not explicitly informed about the word order. We assumed, however, SVO to be the most natural and easiest word order for German native speakers. The noun in subject position always referred to one of the two depicted characters, the noun in object position denoted either the food or the clothing item, depending on the verb's semantic category (food group or clothing group).

Participants' task was to understand the sentences, which, of course, implicitly required learning the meanings of the twelve unknown nouns. They were not explicitly informed about the vocabulary test at the end of the experiment, however. Each of the six inanimate objects and each of the six agent characters was named six times (and each one was shown twelve times) and each of the six verbs was used six times, resulting in 36 trials, which were presented in random order. The combinations of target character, distractor character, target object, and distractor object on the picture (and of the targets in the sentence) were counterbalanced. For each verb, each target character and each target object of the corresponding category (food or clothing) was used twice but each time combined in a different way with each other and with the distractors (e.g., 'clown' and 'sausage' were used together in only one sentence with the verb for 'eat'). The arrangement of characters and objects in the picture was also counterbalanced, such that targets and distractors, as well as characters and objects were in each position (left-right, bottom-top) equally often. The visual salience of characters and objects was counter-balanced automatically: The targets of the food-group were the distractors of the clothing-group and vice versa. There were no semantic associations between any of the characters and objects and characters never faced any of the objects. Eye-movements were monitored in order to get insights into on-line processes during learning, including potential prediction effects.

Vocabulary Test Phase 3 started with a forced-choice vocabulary test (see Figure 3.3). Participants saw four depictions (objects and characters), listened to one spoken noun, and were asked to click onto the picture that matched the noun. Twelve trials were presented, one for each new noun. Combinations of the four options varied but there was always at least one competitor of the same semantic category (character, food, or clothing item, respectively).

Finally, a sentence judgement test was performed. Learners listened to a spoken sentence and their task was to decide whether the sentence was plausible or not. Decisions had to be indicated by pressing a button (YES or NO) on a button box (the YES-button was right for right-handed participants and left for left-handed participants). There were

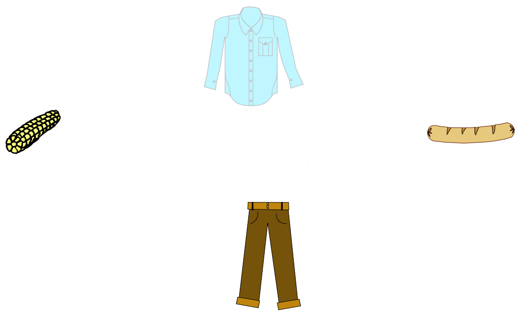


Figure 3.3.: Vocabulary test Exp. 1
si worel
'the sausage'

24 trials, all grammatical SVO sentences which participants had not heard before but which contained only words which were potentially learned. Twelve of the sentences were plausible (e.g., *Si pripas bermamema si worel.*, 'The ballerina will eat the sausage.'), twelve implausible (e.g., *Si worel mankemema si badut.*, 'The sausage will sew the clown.'). Out of the twelve implausible sentences, six had an implausible syntactic subject and six an implausible syntactic object. The implausible-subject sentences either had a food object or a clothing item as agent. In the implausible-object sentences, either a character or an object of the non-matching category (food for clothing-verbs and clothing for food verbs) was the postverbal argument.

3.1.3. Predictions

We first discuss predictions regarding eye-movement behavior. Hypothesizing that participants understand the SVO-sentence structure and show similar gaze behavior as native comprehenders (Wonnacott, 2007), we expected more looks to characters than objects during NP1 and more looks to objects than characters during NP2. We further hypothesized that to identify subject referents and learn their names, participants would exploit CSWL: While in the very beginning of Phase 2 it was not possible for participants to know to which character the noun in NP1 referred, tracking co-occurrences of character names and depicted agents across trials potentially helped them to identify the target. This hypothesis predicts a higher proportion of looks to the target character than to the distractor character to emerge over time. We expected the increase to become visible in two measurements: In the averaged eye-movement data of all trials and of all participants, we expected to see proportionally more looks to the target than to the distractor; in a block analysis (i.e., when dividing the 36 experimental trials in three even blocks), we expected to find most looks to the target in Block 3 (Trials 25-36) and least looks to the target in Block 1 (Trials 1-12). We moreover hypothesized that learners would rapidly exploit verbal restrictions to identify object referents and learn their names (SLCL), possibly additionally to CSWL. That means that during verb and NP2, target objects should be inspected much more often than distractor objects even early in Experiment Phase 2 and there should be no significant development over the course of the experiment (i.e., in the block analysis).

Regarding the forced-choice vocabulary test (Phase 3), our hypothesis that learners apply CSWL and SLCL predicts that learning rates and sentence judgement performance are above chance (25%). Hypothesizing that verbal restrictions provide additional cues regarding target objects (via SLCL) moreover predicts that object names are learned better than character names overall. We moreover made the hypothesis that eye-movements reflect people's successful identification of referents, which potentially results in learning; we therefore expected learning rates to positively correlate with eye-movements. In particular, we predicted the proportion of inspections to the target character during NP1 to correlate with character name learning and the proportion of inspections to the target object during NP2 to correlate with object name learning. Finally, we hypothesized that anticipating referents based on verbal restrictions has a positive effect on object-noun learning and predicted a positive correlation between proportions of inspections to the target object during the verb and object-name learning.

3.1.4. Data Analysis & Results

3.1.4.1. Off-line Results

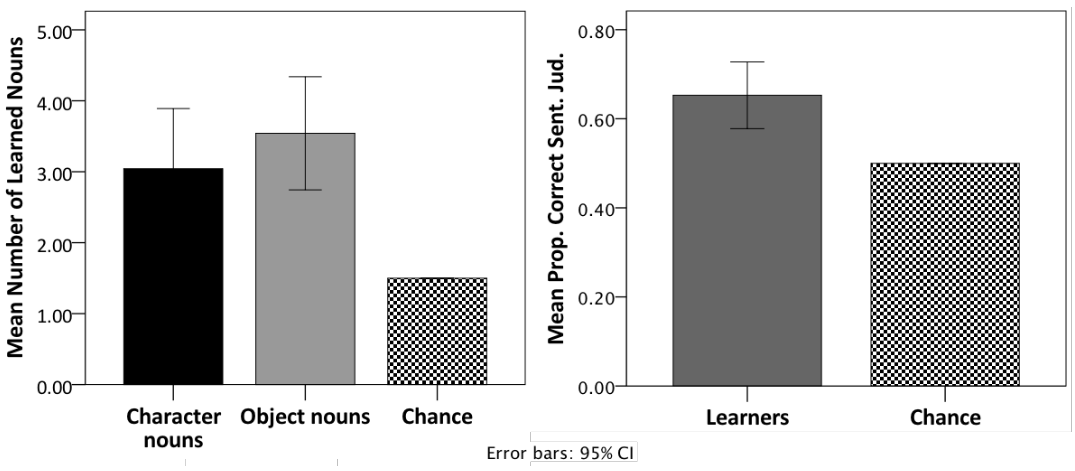


Figure 3.4.: Off-line data Exp. 1: Mean number of learned nouns (left chart) and mean percentage of correct sentence judgements (right chart)

The vocabulary test reveals a word-learning rate of about 55%, which is well above chance (= 25%; $t(23) = 9.276, p < .001$). When analyzing only the data of the participants who showed perfect performance in the verb test for all six verbs (*good verb learners*), $N = 15$, it was 64% ($t(14) = 8.594, p < .001$). There is further a marginally significant positive correlation between verb learning success and noun learning success ($r = .343, p = .05$). There was a tendency for object names to be learned better than character names (all: 59% vs. 51%; only good verb learners: 69% vs. 59%) but this difference is not significant (all: $t(23) = .901, p = .377$; only good verb learners: $t(14) = 1.375, p = .191$). Both object names and character names were learned significantly better than chance (character names: $t(23) = 7.412, p < .001$; object names: $t(23) = 9.191, p < .001$; see Figure 3.4).

The sentence judgement results are significantly above chance (= 50%, possible choices YES or NO) as well: 65% for all participants ($t(23) = 17.443, p < .001$) and 73% for good verb learners ($t(14) = 16.977, p < .001$). Results of the sentence judgements are positively correlated with the vocabulary learning results ($r = .449, p = .014$) and with the verb learning results ($r = .411, p = .023$).

3.1.4.2. On-line Results

For eye-movement analysis, we examined inspections on each region of interest (ROI; target character, distractor character, target object, distractor object) for three time periods linked to the unfolding sentence (from onset of NP1 to onset of verb (V): 1200-3100ms, from onset of V to onset of NP2: 3100-5100ms, and from onset of NP2 to offset of NP2: 5100-6500ms). Specifically, we measured for each of the four ROIs in each of the three time periods whether there was at least one inspection in a trial. The analyzed NP1 region started 200ms later than the actual NP1 region because in that time the non-differentiating determiner was heard. The same was done for NP2, which means that verb region ended 200ms later.

We conducted logistic regression analyses by entering the binomial data (inspection or no inspection at certain time to a specific ROI) into linear mixed effects (*lmer*) models with a logit link function (from the lme4 package in R, Bates, 2005). Participants and items were considered as random factors. To see whether the fixed factor (ROIs) had a main effect (i.e., whether including the factor significantly improved the predictive power of the model, regarding where people looked) we compared the models that include and exclude this factor with a chi-Square test (Baayen, Davidson & Bates, 2008). Contrasts between levels of a factor (single ROIs) were assessed by the ratio of regression coefficients and standard errors since the p-values produced by lmers (Wald z test) are anti-conservative (Baayen et al., 2008): If the coefficient was larger than twice the standard error, the difference was considered to be significant. Tables of these statistical comparisons are provided below.

Table 3.1.: Lmer models for inspections on characters vs. objects and targets vs. distractors during time periods, Exp. 1

$InspectionsduringNP1/V/NP2 \sim ROI + (1|sub) + (1|item), family = binomial(link = "logit")$

		Predictor	Coef.	SE	Wald z	p
characters vs. objects						
1	*NP1*	(Int) (char)	0.822	0.070	11.730	$< .001$
2		objects	−0.992	0.074	−13.360	$< .001$
3	*V*	(Int) (char)	−0.300	0.085	−3.532	$< .100$
4		objects	0.681	0.073	9.365	$< .001$
5	*NP2*	(Int) (char)	−0.871	0.093	−9.357	$< .001$
6		objects	0.453	0.076	6.00	$< .001$
target vs. distractors						
7	*NP1*	(Int) (targ)	0.836	0.144	5.810	$< .001$
8		distractor	0.027	0.111	0.238	$= .812$
9	*V*	(Int) (targ)	0.477	0.123	3.886	$< .001$
10		distractor	−0.179	0.104	−1.716	$< .010$
11	*NP2*	(Int) (targ)	−0.212	0.100	−2.134	$< .050$
12		distractor	−0.406	0.104	−3.912	$< .001$

The formulas describing the lmer models are of the following form: dependent variable (inspections during time period) is a function of (∼) the independent variable (ROI) plus random effects (subjects and items). Analyses were conducted using the R statistical package.[1]

There were reliably more inspections on the characters than on the objects during NP1 (Table 3.1, Rows 1-2) and reliably more inspections on the objects than on the characters in the V interval (Rows 3-4) and NP2 (Rows 5-6). The difference between looks to target character and distractor character in NP1 was not significant (Rows 7-8). However, there was a significant difference in V: The target object was inspected more than the distractor object (Rows 9-10). Moreover, the target object was looked at significantly more than the distractor object during NP2 (Rows 11-12; see Figure 3.5).

For illustration, we plotted inspections to the four regions of interest over time (Figure 3.6). Time-region borders correspond to the ones used for inferential analyses (onset noun of NP1 until onset verb, onset verb until onset noun of NP2, and onset noun of NP2 until end of the recording), averaged across items. The graph shows that during the noun of NP1 (from 1200ms to about 3100ms), inspections to the characters (dotted lines) rise while objects (full lines) are looked at less. In the verb region (3100ms to about 5100ms),

[1]www.r-project.org

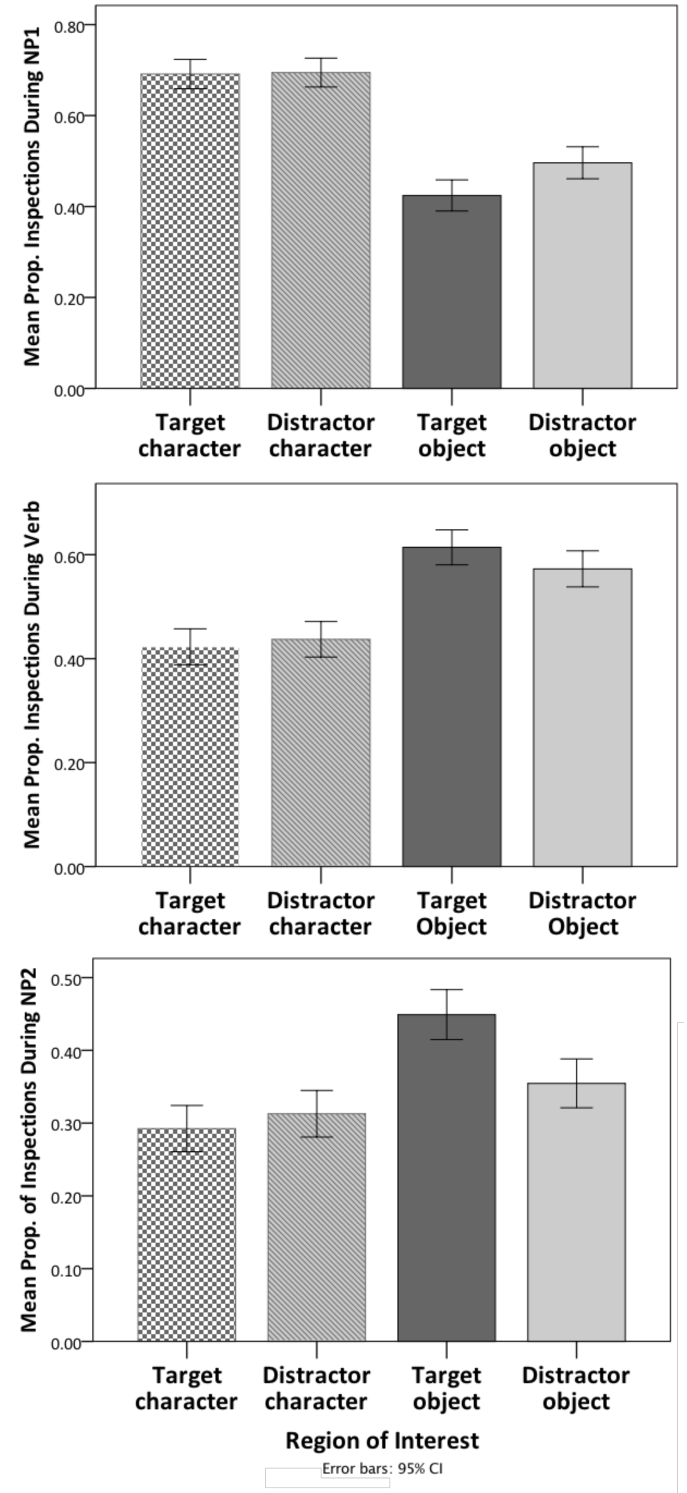


Figure 3.5.: Proportions of trials with at least one inspection to ROIs during NP1 (top chart), the verb (mid chart), and NP2 (bottom chart), Exp. 1

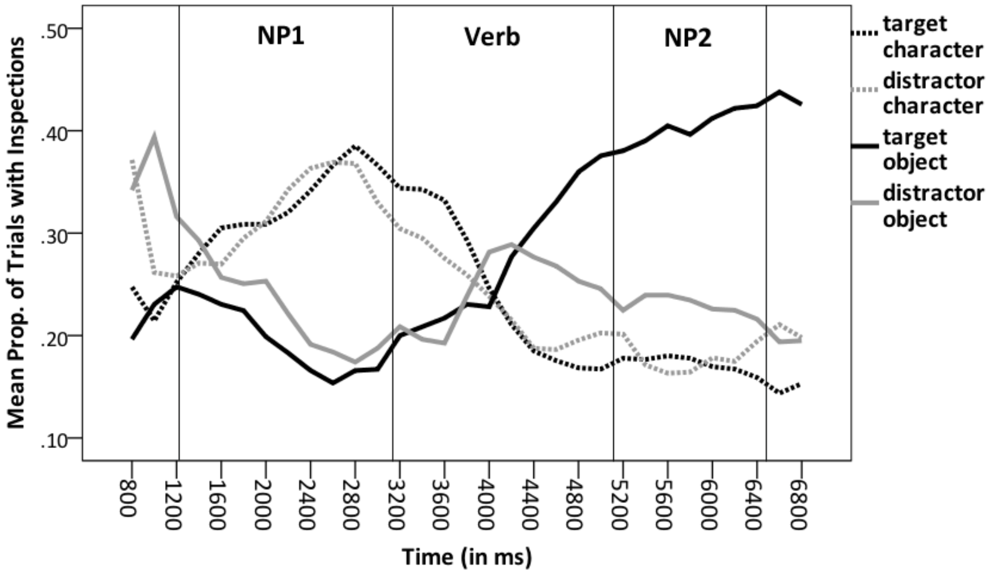


Figure 3.6.: Timegraph Exp. 1

inspections to the objects rise while inspections to the characters decrease. This trend continues during NP2 (5100ms to about 6500ms). There is also a clear increase in looks to the target object and a decrease in looks to the distractor object starting at around 4200ms, during verb region.

To investigate potential learning development over the course of the experiment, we divided trials (in order of presentation) into four blocks. We entered the proportions of inspections to the target character during NP1 in Blocks 1 to 4 into our analysis. No main effect was found ($\chi(3) = 3.504, p = .320$; see Figure 3.7). Likewise, the proportions of inspections to the target object during the verb and during NP2 were compared. Again, we found no main effects and differences between blocks were minor. (V: $\chi(3) = 2.235, p = .525$; NP2: $\chi(3) = 3.302, p = .347$, Figure 3.7).

We further examined the relation between inspections and learning performance. We found a significant positive correlation between proportions of inspections to the target object during NP2 and object-name learning ($r = .648, p < .001$). The positive correlation between learning performance for character names and proportions of inspections on the target character during NP1 was only marginally significant ($r = .080, p = .081$). Finally, no significant correlation between inspecting the target object during the verb (again, as proportions) and object-name learning was found ($r = .218, p = .150$).

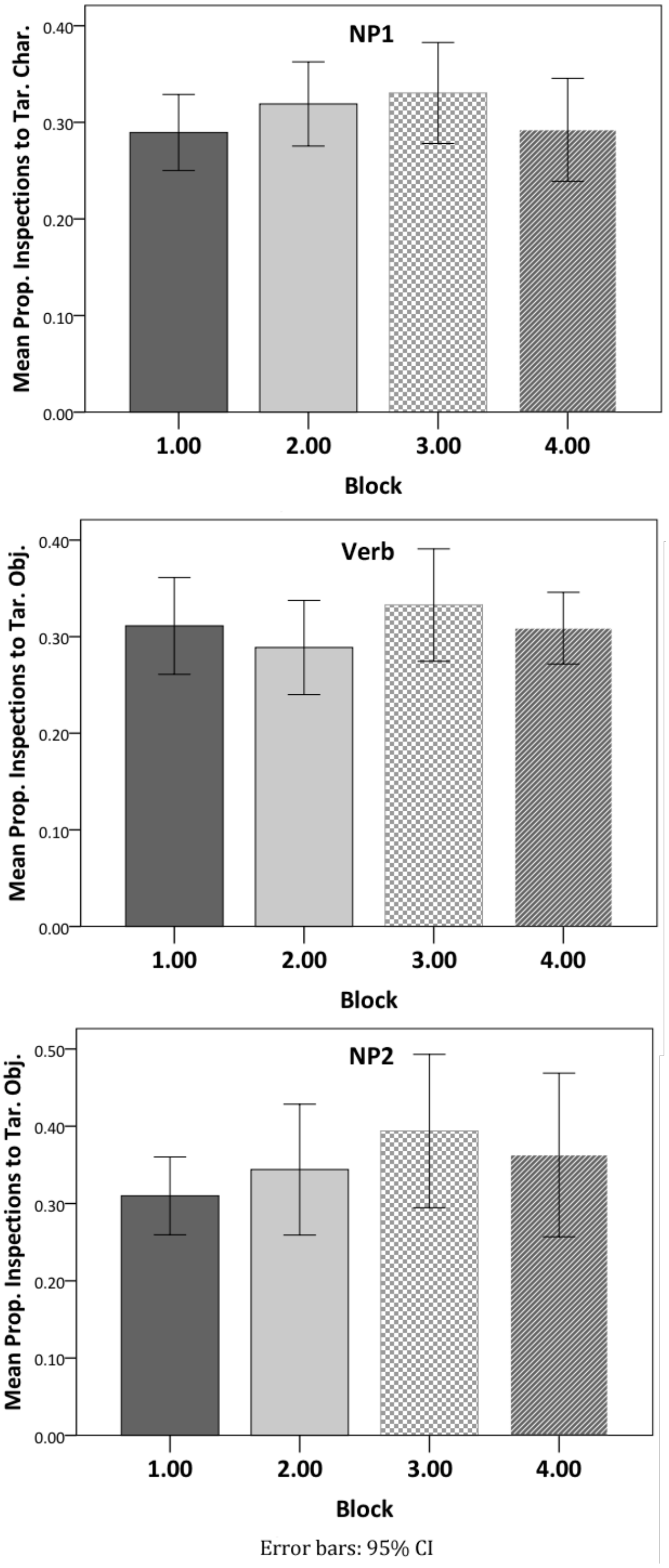


Figure 3.7.: Proportions of trials with at least one inspection to target character during NP1 (top chart), the verb (mid chart), and NP2 (bottom chart), Exp. 1

3.1.5. Summary and Discussion

Data from Experiment 1 reveals that while there was a tendency for object-name learning to be better than subject-name learning, both learning rates were significantly above chance. Sentence-judgement results highly correlate with learning rates. Good verb learners performed better than bad verb learners in both tasks. Eye-movements reflect that participants preferred to inspect the agent characters during NP1 and the inanimate objects during the verb and NP2. There was moreover a tendency to look more at the target character than the distractor character during NP1. During the verb and NP2, learners looked more often at the target object than at the distractor object. This difference was significant in NP2 and marginally significant in the verb region. We further found positive correlations between character-name learning and looks to the target character during NP1 (marginal effect) as well as between object-name learning and looks to the target object during NP2 (significant effect). No effects were found when comparing inspections to target regions (target character during NP1 and target object during V and NP2) between blocks (Block 1: Trials 1-12, Block 2: Trials 13-24, Block 3: Trials 25-36).

As expected, participants were successful in learning nouns, both names of characters and names of objects. While character-name learning suggests that CSWL is indeed a useful mechanisms for learners in the naturalized situation we used (i.e., with nouns as parts of sentences and visual referents embedded in scenes), object-name learning most likely resulted from SLCL (verbal constraints) unambiguously identifying referents. Theoretically, object-name learning could be also based on CSWL but we take this interpretation as very unlikely given eye-movement data: The difference between looks to the target object and the distractor object during NP2 is very clear and much more pronounced than the difference between looks to the target character and the distractor character during NP1. Moreover, the fact that the target object was looked at significantly more often than the distractor object during the verb suggests that verb information influenced object-name learning. We further observed that good verb learners performed better in noun learning. We take this observation into account for the development of the following experiments. However, we cannot draw conclusions about causality: While verb learning may have helped noun learning (which would be plausible given the verbal constraints), it is also possible that good verb learners simply tended to be good noun learners.

Learners' gaze pattern was generally native-like: They inspected the characters more than the objects during NP1 and the objects more than the characters while NP2 was spoken. Eye-movements additionally reflected word learning given that targets were inspected more often than distractors. Importantly, this difference is much larger during NP2 than during NP1, clearly revealing the expected boost of object target identification caused by verbal constraints (SLCL). However, the finding that the difference between inspections to the target character and to the distractor character during NP1 was not even significant is not as expected. This lack of effect may have been caused by a ceiling effect due to animacy: People generally tend to attend to animates more than inanimates (Langton, Law, Burton & Schweinberger, 2008; Boland, 2004), even if they are only drawn (Sagiv & Bentin, 2001; Tomalski, Csibri & Johnson, 2009). This may have caused a high number of automatic looks to both characters, potentially blocking the difference between looks to the target and looks to the distractor which was based on learning from becoming reliable.

As mentioned above, gaze also reflects that participants anticipated the target object during the verb already, providing evidence for a very rapid effect of SLCL on on-line identification of referents. We cannot, however, conclude that anticipation of the target object had a positive effect on object-noun learning given that the positive correlation between inspecting the target object during the verb and learning object names did not reach significance. However, we suggest that the effect may have been present but too subtle to be reflected by this measurement.

While not statistically reliable, the tendency for objects to be learned better than characters, may be taken as preliminary evidence that SLCL not only boosted learners' on-line identification of referents but additionally bootstraps noun learning. We suspect that the animacy effect noted above may have kept this difference from becoming more pronounced: It may have caused a boost for character name learning in this experiment which equalized the boost of object name learning based on SLCL.

The close relationship between on-line interpretation and final learning is further supported by the positive correlations between character-name learning and looks to the target character as well as between object-name learning and looks to the target object. While the first of these two correlations was only marginally significant, we attribute this weakness, again, to ceiling effects.

Finally, block analysis of gaze did not reveal any developmental effects in inspections to target regions (target character during NP1 and target object during V and NP2). While this lack of effect was expected for the verb region and NP2 because we predicted object-referent identification to be successful from early on, it is surprising for NP1. However, it is again possible that a ceiling effect blocked a difference to clearly emerge: Firstly, CSWL was still relatively easy in this experiment and, secondly, all agent characters were visually salient (partly due to their animacy) and therefore looked at frequently.

To summarize, Experiment 1, firstly, provides evidence for the claim that CSWL and (most likely) SLCL are useful word-learning mechanisms in the naturalized situations we utilized. Secondly, we revealed that learners at a very early stage can already follow a sentence in a native-like way: Participants incrementally and predictively integrated linguistic input and visual cues into their comprehension as reflected by eye-movements.

Looks to the target object during verb and NP2 moreover clearly support the hypothesis that SLCL boosts language novices' on-line understanding. We cannot, however, draw clear conclusions about a possible bootstrapping effect on actual noun learning, given the non-significance of the comparison between character-name and object-name learning. This question will be addressed in Experiment 3 again. The role of anticipating object referents for noun learning also remains unclear given that we did not find a correlation between anticipating referents and learning referents' names. Experiment 2 sought to further investigate this issue.

3.2. Experiment 2

3.2.1. Motivation

The aims of Experiment 2 were, firstly, to replicate and generalize the results we revealed in Experiment 1 and, secondly, to explore the possible role of the verb-driven anticipatory effects on noun learning success. To accomplish these two objectives, a second word order was introduced as between-participant condition: object-verb-subject word order (OVS). The purpose of this manipulation was to compare learning behavior when participants can anticipate the syntactic object when following verb-driven expectations (SVO) and when they cannot because they receive the verb information following the syntactic object (OVS). We hypothesized that learning of object nouns would be better in SVO than OVS. Otherwise the design was similar to Experiment 1.

3.2.2. Methods

3.2.2.1. Participants

44 German native speakers participated in Experiment 2 for a reimbursement of € 5, four of which had to be excluded due to technical problems and experimenter error. Resulting data of 40 participants was included in the analyses (10 male, 30 female, aged between 19 and 46, average age 25, mainly students from various disciplines). None of the participants had participated in Experiment 1.

3.2.2.2. Materials and Procedure

Materials of Experiment 2 were almost identical to those in Experiment 1 (see Section 3.1.2.2), with the following changes. Firstly, verb phonology was simplified to possibly facilitate verb learning (the verb suffix was reduced to -ma and some verb stems were changed). Moreover, some nouns were phonologically changed in order to make them sound more distinct. There were two mapping lists again. Secondly, for 20 of the 40 participants, the word order in Phase 2 was OVS instead of SVO.

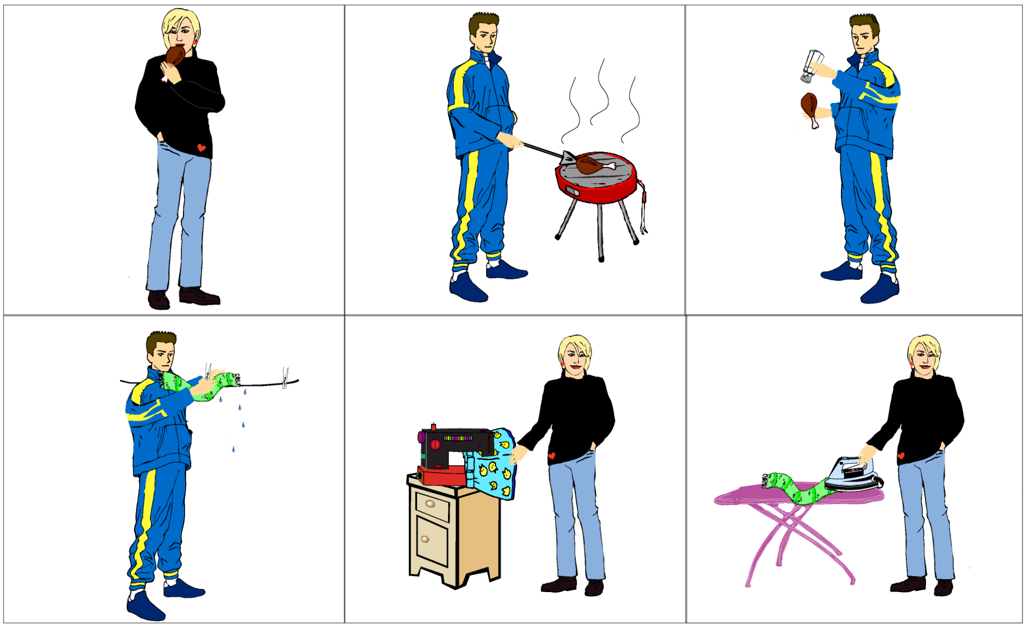


Figure 3.8.: Verb Learning, Exp. 2

The procedure of Experiment 2 was similar to the procedure of Experiment 1, as well. Again, the experiment consisted of two learning phases and a testing phase. The verb learning procedure (Phase 1) was changed from that in Experiment 1, with the aim of enhancing learning success: Pictures of the six actions were presented simultaneously and participants were asked to click onto them to play the action names self-paced. They were encouraged to click onto the different actions as often as they wanted and in their preferred pace and order, for approximately 5 minutes. The verb testing phase was as in Experiment 1 (see Figure 3.8). The eye-tracker was adjusted and verbs were shortly repeated before the noun-learning phase started. Instructions were more explicit in terms of word order because it was important to make people aware of the unusual word order in OVS, in order to avoid a much longer orientation phase compared to SVO. People were prompted, again, to understand the sentences but it was also mentioned in the instructions that this required learning the unknown words. The procedure for Phase 2 was the same as in Experiment 1, as was the forced-choice vocabulary test. The experiment lasted approximately 30 minutes.

3.2.3. Predictions

We predicted that object-name learning would be better in Condition SVO than in Condition OVS, hypothesizing that verb-driven anticipation, which is possible in SVO but not in OVS, boosts noun learning. We should therefore see an interaction between the two factors

Word Order (SVO vs. OVS) and Noun Type (character name vs. object name): While character-name learning should be similar in both conditions, object-name learning should be better in SVO than OVS. Assuming that we improved the verb learning procedure, we further expected a higher noun learning rate in SVO than in Experiment 1.

The pattern of eye-movements in SVO was predicted to be as in Experiment 1: We expected a preference for looks to the characters during NP1 and a preference for looks to the objects during NP2, a preference for looks to the targets during both NP1 and NP2, and anticipatory inspections to the target object during verb. For OVS, on the contrary, we expected a different pattern: More looks to the objects than to the characters during NP1, a preference for looks to the target object during verb, and more looks to the characters than to the objects during NP2.

3.2.4. Data Analysis & Results

3.2.4.1. Off-line Results

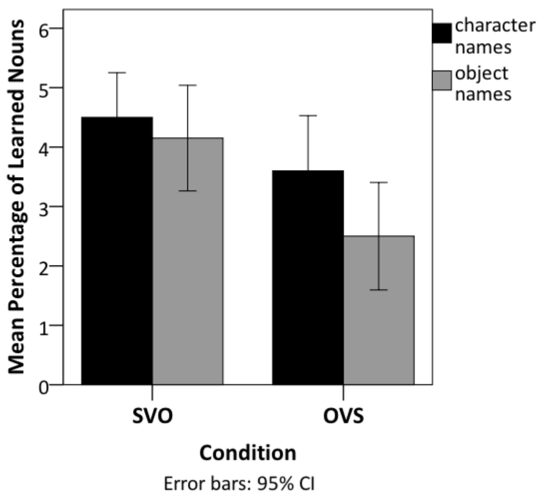


Figure 3.9.: Mean number of learned nouns, Exp. 2

In SVO, participants learned 72% of the twelve nouns, which is significantly above chance ($t = 8.249, p < .001$) and a clear improvement over the first experiment. Participants in OVS performed significantly worse ($t(38) = 2.466, p < .05$): They learned only 51% of the nouns (still reliably above chance, 25%: $t = 3.840, p < .001$). While we found a significant difference between object-name learning in SVO and OVS ($t(38) = 2.723, p < .05$), there was no significant difference between character-name learning in SVO and OVS ($t(38) = 1.577, p = .123$), see Figure 3.9. However, the interaction between factors Word Order (SVO vs. OVS) and Noun Type (character name vs. object name) did not reach significance ($\chi(1) = 2.038, p = .153$).

Table 3.2.: Lmer models for inspections on characters vs. objects and targets vs. distractors during time periods, Exp. 2, SVO

$$InspectionsduringNP1/V/NP2 \sim ROI + (1|sub) + (1|item), family = binomial(link = "logit")$$

		Predictor	Coef.	SE	Wald z	p
characters vs. objects						
1	*NP1*	(Int) (char)	1.864	0.151	12.370	$< .001$
2		objects	−1.896	0.135	−14.010	$< .001$
3	*V*	(Int) (char)	0.334	0.142	2.353	$< .050$
4		objects	0.847	0.120	7.084	$< .001$
5	*NP2*	(Int) (char)	−0.479	0.141	−3.402	$< .001$
6		objects	1.520	0.119	12.747	$< .001$
target vs. distractors						
7	*NP1*	(Int) (targ)	0.795	0.142	5.608	$< .001$
8		distractor	−0.142	0.119	−1.196	$= .232$
9	*V*	(Int) (targ)	0.081	0.130	0.627	$= .530$
10		distractor	0.159	0.116	1.370	$= .171$
11	*NP2*	(Int) (targ)	0.423	0.100	4.244	$< .001$
12		distractor	−0.784	0.116	−6.756	$< .001$

3.2.4.2. On-line Results

Eye-movements were analyzed as in Experiment 1 (see Section 3.1.4.2). Analysis for SVO revealed the same patterns as in Experiment 1 (see Table 3.2 and Figure 3.10): During NP1, participants inspected the characters significantly more often than the objects (Table 3.2, Rows 1-2), and the target non-significantly more often than the distractor (Rows 7-8). During the verb, objects were inspected significantly more often than characters (Rows 3-4), and the target non-significantly more than distractors (Rows 9-10). During NP2, participants inspected objects significantly more than characters (Rows 5-6) and the target significantly more than the distractor (Rows 11-12).

Eye-movements in OVS show a different pattern (see Table 3.3 and Figure 3.10): During NP1, characters are inspected significantly more often than objects (Rows 1-2) but there is no significant difference between looks to the target character and the distractor character (Rows 7-8). When the verb is spoken, objects are inspected significantly more than characters (Rows 3-4) and the target object is looked at non-significantly more than the distractor object (Rows 9-10). During NP2 region, characters are inspected reliably more often than objects again (Rows 5-6). The target character is looked at most but the difference to the distractor character is not significant (Rows 11-12). See graphs for both conditions and all time regions in Figure 3.11.

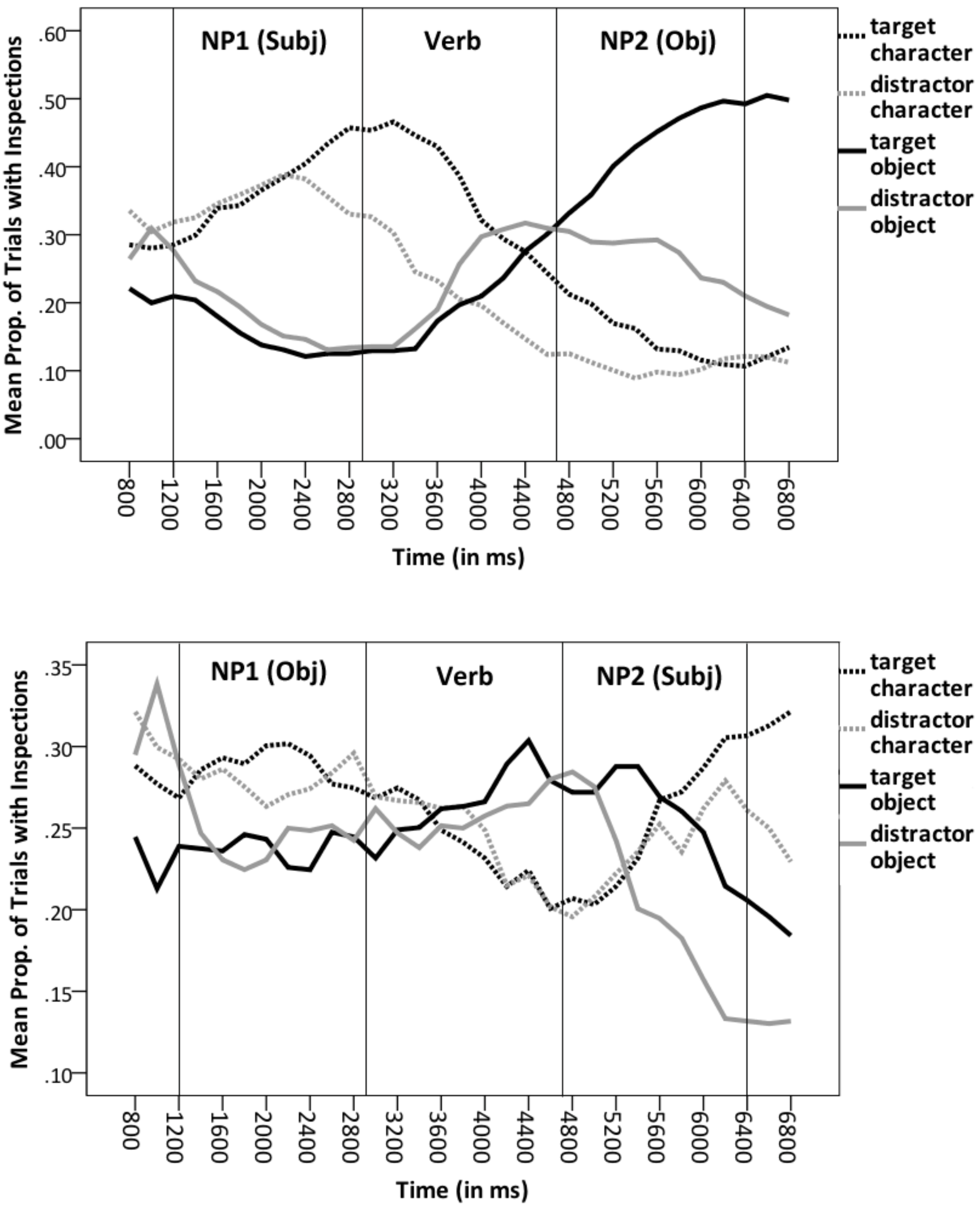


Figure 3.10.: Timegraph Exp. 2, SVO (top) and OVS (bottom)

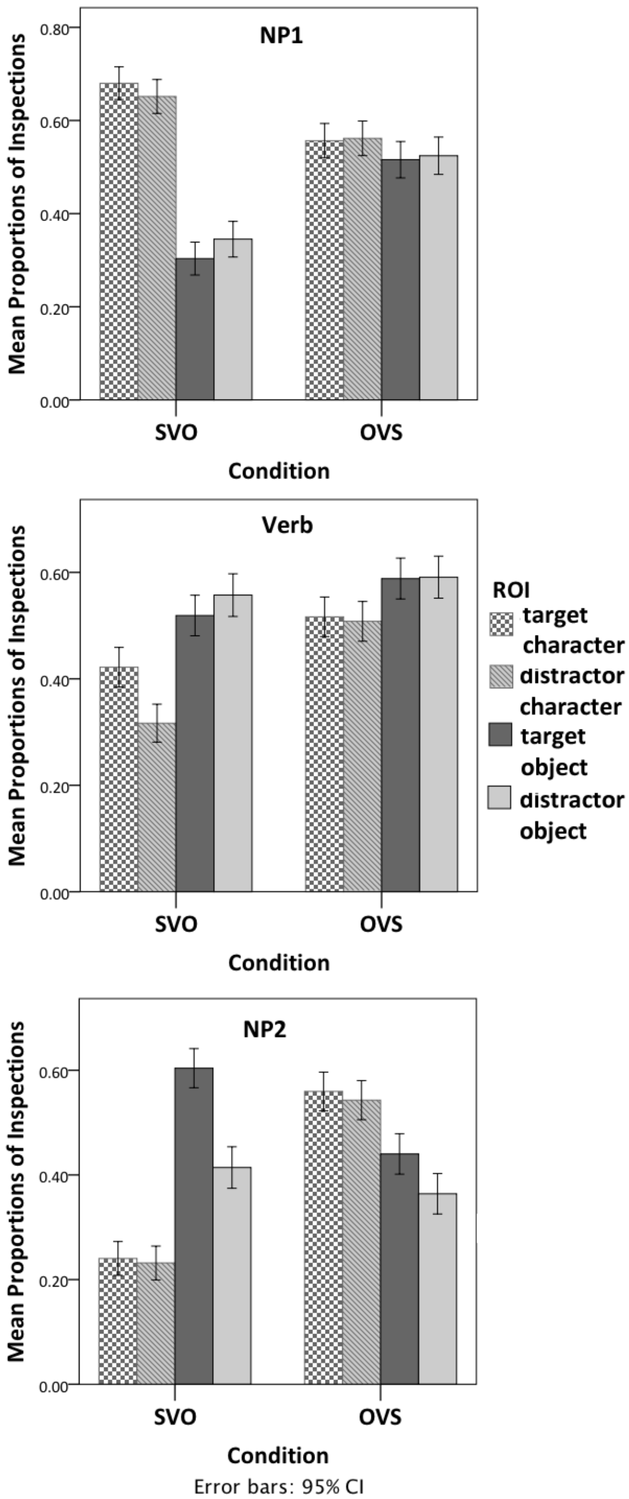


Figure 3.11.: Proportions of trials with at least one inspection to ROIs during NP1 (top chart), the verb (mid chart), and NP2 (bottom chart), Exp. 2

Table 3.3.: Lmer models for inspections on characters vs. objects and targets vs. distractors during time periods, Exp. 2, OVS

$InspectionsduringNP1/V/NP2 \sim ROI + (1|sub) + (1|item), family = binomial(link = "logit")$

		Predictor	Coef.	SE	Wald z	p
characters vs. objects						
1	*NP1*	(Int) (char)	1.324	0.158	8.372	$< .001$
2		objects	-0.367	0.128	-2.869	$< .010$
3	*V*	(Int) (char)	1.194	0.153	7.809	$< .001$
4		objects	0.332	0.134	2.482	$< .050$
5	*NP2*	(Int) (char)	1.180	0.150	7.884	$< .001$
6		objects	-0.745	0.121	-6.140	$< .001$
target vs. distractors						
7	*NP1*	(Int) (targ)	0.227	0.162	1.402	$= .161$
8		distractor	0.013	0.112	0.116	$= .907$
9	*V*	(Int) (targ)	0.369	0.117	3.152	$< .010$
10		distractor	0.009	0.117	0.074	$= .941$
11	*NP2*	(Int) (targ)	.241	0.189	1.273	$= .203$
12		distractor	-0.074	0.114	-0.649	$= .516$

Based on this eye-movement data, we additionally selected a subset of OVS-participants who made more looks to the two inanimate objects than to the two animate characters during NP1 and more looks to the characters than to the objects during NP2 (N=7, see Figure 3.12). We found that this group had a numerically higher learning rate than the group of non-selected OVS-participants (58%, $t(18) = -0.832, p = .417$). While none of the differences between SVO and OVS-selected in noun learning, object-name learning, and character-name learning were significant (nouns: $t(25) = 1.177, p = .250$; objects: $t(25) = 0.640, p = .180$; characters: $t(25) = 1.177, p = .530$), the difference in object-name learning was larger (19%) than the difference in character-name learning (8%). The interaction between factors Word Order and Noun Type, however, was not significant ($\chi(2) = 2.232, p = .323$).

3.2.5. Summary and Discussion

Data from Experiment 2 reveals that learning rates in both SVO and OVS were significantly above chance but significantly better in SVO than OVS. While learning rates for object names were significantly higher in SVO than OVS, the difference between performance in both conditions concerning character-name learning was not significant. The interaction between factors Word Order and Noun Type did not reach significance, however. Learners

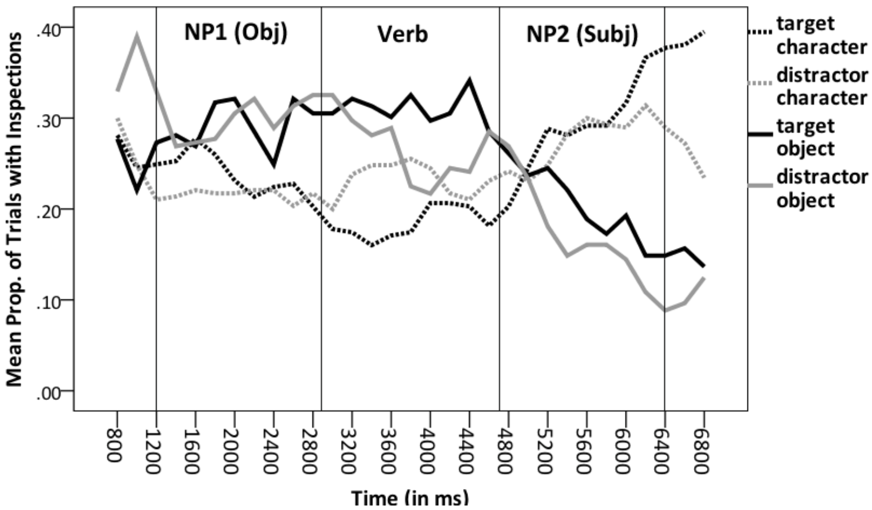


Figure 3.12.: Timegraph Exp. 2 - OVS selected

belonging to the OVS-selected group performed numerically better in noun learning than the other OVS-participants. While there was no significant interaction, the difference in learning between conditions was more pronounced for object names than for subject names. Eye-movements in SVO reflect the same pattern as eye-movements in Experiment 1: During NP1, agent characters were looked at significantly more often than inanimate objects, during the verb and NP2 it was the other way around. Moreover, targets were inspected more often than distractors but this effect was significant only for region NP2. In Condition OVS, characters were inspected significantly more often than objects during both NP1 and NP2; the opposite pattern was found for the verb region.

On-line and off-line results in Condition SVO pattern as expected and replicate the findings from Experiment 1, with an overall improved word learning rate, possibly due to the revised verb-learning procedure (Phase 1). The only drawback here is the non-significance of the difference between looks to target object and distractor object during verb region. This, however, is most likely due to the shortening of the verb phase - the suffix was only one syllable long, which might have been too short for an effect to show up clearly enough in this region.

Noun learning in OVS was also successful. Eye-movements seem to reflect a combination of both following the sentence structure and searching for referents, at least during verb and NP2: Shortly after hearing the verb, learners' looks to the objects increased. While this is plausible because the verb disambiguated which of the two depicted objects was the referent of the object noun, we did not find a significant difference between looks to both objects. Immediately after the onset of NP2 (that is, even before the noun was named, see timegraph in Figure 3.10), looks to the animate characters increased which probably reflects following of the sentence structure. The eye-movements during NP1 (more looks to the characters than to the objects), however, are surprising. One potential interpretation is that participants may not have understood the word order, possibly because German native speakers are very much used to subject-before-object word order in simple sentences. However, it is also possible that participants did not look at the objects during NP1 because, firstly, they were waiting for the verb information to disambiguate between both objects and, secondly, the characters were much more salient than the objects (see discussion about animacy in Experiment 1). The fact that the learning rate in OVS is clearly above chance and in fact very similar to the learning rate in Experiment 1 strongly supports the assumption that learners were generally capable of processing the sentences. It is also possible that learners entirely ignored verbs and the sentence structure and only attended to the nouns to conduct CSWL for both nouns. However, we think that this explanation is unlikely because eye-movements during verb and NP2 do reflect that learners considered both the verb and the sentence structure.

We found some interesting differences in learning rates between conditions. Firstly, learning of all nouns was reliably worse in OVS than SVO. Secondly, this difference was more pronounced in object-name learning than character name learning (even if the interaction was not significant). This is broadly in line with our prediction that object noun learning is easier when the object follows the verb (and can be anticipated), than when it precedes the verb. However, given the uncertainty about the eye-movement pattern in OVS, we must conclude carefully. The fact that the object-name learning rate in the OVS selection is still worse than the one in SVO indicates that even if eye-movements more clearly reflect an understanding of the OVS word order, it might be more difficult to learn object names (the difference is non-significant, however, this could be because N is very small). However, even successful OVS-comprehension during the experiment is unlikely to offset the natural difference in ease that learners have with the two word orders and we still cannot clearly differentiate between a potential influence of the relative position of the verb, on one hand, and a potential influence of the absolute sentence-initial position on other hand.

To summarize, data from Experiment 2 supports the view that CSWL, firstly, is a useful mechanism in non-instructed situations, secondly, works with different word orders, and, thirdly, can lead to high learning rates (SVO). Our results are furthermore at the least consistent with the hypothesis that verb-driven anticipation of a referent boosts noun learning. Finally, eye-movement data in OVS suggests that learners were influenced by an interesting combination of processes. We now further investigate the combined influence of SLCL and CSWL when both are applicable in a complementary way in Experiment 3.

3.3. Experiment 3

3.3.1. Motivation

The results of Experiment 1 were not entirely conclusive with regard to the influence of verbal restrictions (SLCL) on object-name learning (see discussion in Section 3.1.5). While we found a very clear advantage in the on-line identification of the referents denoted by the syntactic object compared to referents denoted by the syntactic subject, there was only a tendency for improved learning of object nouns compared to CSWL-learned subject nouns. One potential confound which may have masked a significant effect on learning rates is the animacy of the subject referents. Possibly, participants attended more to the animate agents than to the inanimate objects, facilitating subject noun learning.

Addressing this animacy issue in Experiment 3, we based our observations entirely on inanimate object referents and manipulated the degree of verbal restriction to study the direct influence of verbal constraints (SLCL) on noun learning. Our hypothesis was that noun learning would be better if SLCL was available than noun learning which is based on CSWL alone. Additionally, we investigated the interaction of CSWL and SLCL when they were complementary not only concerning the learning of a set of nouns, as in Experiments 1 and 2, but with regard to the learning of a single noun. In particular, we included a condition where SLCL was applicable but weak by leaving a small degree of referential uncertainty. We hypothesized that to compensate for this uncertainty, CSWL would be applied in addition to SLCL. Our hypotheses were based on the assumption that SLCL would be used in a direct and deterministic way whenever it is available, as supported by eye-movements in Experiment 1: During the verb and NP2, verbal constraints led to an immediate and unambiguous identification of the correct object referents while all other objects were not considered.

Specifically, we compared object noun learning in the following three conditions: In the first condition (*No Referential Uncertainty*, *NoRU*), the restrictive verb (SLCL) identified exactly one referent in the scene. In the second condition *Low Referential Uncertainty*, *LowRU*), the restrictive verb (SLCL) identified two potential referents in the scene. In the third condition (*High Referential Uncertainty*, *HighRU*), the non-restrictive verb did not constrain the semantic category of the referent which means that there were four potential referents. We hypothesized that noun learning in NoRU and LowRU would be easier than noun learning in HighRU, due to SLCL decreasing referential uncertainty (from four to one in NoRU and from four to two in LowRU). Furthermore, we hypothesized that, in LowRU, learners would apply CSWL additionally to SLCL in order to compensate for the remaining level of uncertainty. We hypothesized that learning would be either worse in LowRU than NoRU or equally good in both conditions. The first possibility would be supported by the facts that, firstly, referential uncertainty in LowRU was higher than in NoRU and, secondly, the additional use of CSWL in LowRU may require more cognitive capacities, reducing the overall performance. The rationale behind the second hypothesis is that the complementary use of CSWL in LowRU may completely compensate for the difference in referential uncertainty between NoRu and LowRU: Applying both SLCL and CSWL may result in greater engagement by the learner.

3.3.2. Methods

3.3.2.1. Participants

50 native speakers of German, most of them students (from different disciplines) participated in Experiment 3 (23 female, 9 male, aged between 19 and 31, average age 24). Participants, who received €5, had not participated in Experiments 1 or 2. Given that 18 learners had to be excluded due to bad verb learning or technical problems, data of 32 participants entered analyses.

3.3.2.2. Materials and Procedure

The language comprised six verbs, 14 nouns, and the article 'si'. There were two non-restrictive verbs ('take' and 'point at') and four restrictive verbs: depending on experiment list either two food verbs ('eat' and 'barbecue') and two clothing verbs ('iron' and 'sew') (Lists 1, 2, 5, 6) or the two food verbs and two container verbs ('fill' and 'empty') (Lists 3, 4, 7, 8). Nouns denoted two characters (man and woman) and twelve objects: four food items ('broccoli', 'chicken', 'tomato', 'toast'), four clothing items ('trousers', 'tshirt', 'skirt', 'jacket'), and four container items ('vase', 'bottle', 'watering can', 'bucket'). Word order was SVO.

The experiment consisted of five parts with similar procedures as in Experiments 1 and 2. The main difference was that instead of one noun-learning phase and one vocabulary-testing part, there were two each (Blocks 1 and 2, see explanation below). The whole experimental sequence comprised: verb learning and testing, eye-tracker preparation, and verb repetition (Phase 1); noun-learning Block 1 (Phase 2); vocabulary-test Block 1 (and verb repetition) (Phase 3); noun-learning Block 2 (Phase 4); vocabulary-test Block 2 (Phase 5). The experiment lasted approximately 30 minutes.

Verb Learning Phase 1 resembled Phase 1 in Experiment 1 (see Paragraph *Verb Learning* in Section 3.1.2), except that we used animated verb-learning and verb-testing pictures to improve recognizability of the actions. Per animation, three pictures were used. Each verb was repeated eight times with two different animations each (different combinations of a man or a woman manipulating different objects). The eye-tracker was prepared as in Experiment 1 (see Paragraph *Verb Learning* in Section 3.1.2).

Noun Learning In Phases 2 and 4, scene-sentences pairs were presented. In contrast to Experiments 1 and 2, we switched from circus situations to more neutral ones (see Figures 3.13). Items were manipulated according to one three-level within-participant factor (degree of referential uncertainty). There were three conditions: the no-referential-uncertainty condition (Condition NoRU), the low-referential-uncertainty condition (Condition LowRU), and the high-referential-uncertainty condition (Condition HighRU). Firstly, the conditions differed concerning verb type: In Conditions NoRU and LowRU, a restrictive verb was used, in Condition HighRU, it was a non-restrictive verb. Secondly, there were differences in the visual scenes. Images always depicted one character and four objects embedded in a simple indoor scene. One of the objects was the target object. The others were competitors (potential referents matching the verb's constraints) and distractors (objects which were not potential referents). The combination of competitors and distractors depended on the condition the item was in (and the list the participant belonged to): In Condition NoRU, there was no competitor since the verb was restrictive (e.g., 'eat') and only the depicted target fulfilled the verbal constraints (e.g., food). In Condition LowRU, there was one competitor: The verb was restrictive but there was one depicted object in addition to the target which was a member of the verb's required semantic class. That means that while conducting SLCL was possible in both Conditions NoRU and LowRU, it was only sufficient on Condition NoRU. In Condition HighRU, there were three competitors because the verb was non-restrictive and did not semantically constrain the category of potential referents denoted by the post-verbal argument (see Table 3.4).

In Block 1 (Phase 2), no target was a competitor in another trial to make sure participants could not exclude competitors based on other learned words. In Block 2 (Phase 4), however, learning was potentially simplified by the possibility to exclude already learned mappings. The order of presentation of the 48 trials (24 per Block) was randomized, with each noun repeated four times. There were eight lists that accounted for counterbalancing issues and potential confounds: Firstly, as in Experiments 1 and 2, there were two different world-word mappings to avoid that learning effects could be subject to particular associations between words and referents. Secondly, we used two different assignments for objects to conditions (e.g, in half of the lists food objects were in Condition NoRU and for the other half they were in Condition LowRU). Finally, there were two versions of lists that depended on the assignment of items to Blocks 1 and 2. Pictures were counterbalanced in the same way as in Experiments 1 and 2 except that we additionally controlled the objects' spatial relation to the character in the scene to avoid artifacts due to the distance between agents and objects (participants could, for instance, be biased to believe the object close to the agent must be

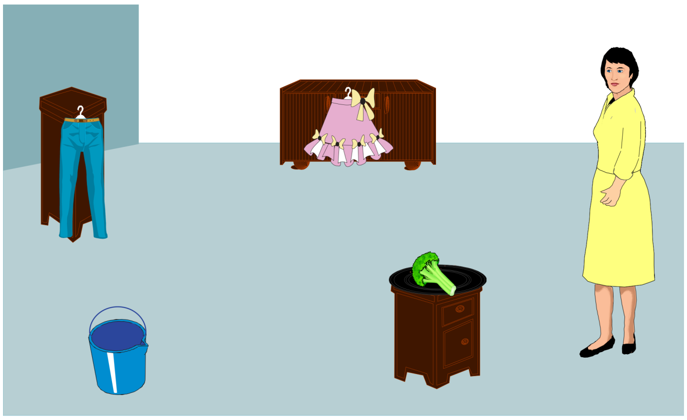


Figure 3.13.: Example item Exp. 3
NoRU: *Si gadis mautimema si sonis.* ('The woman will empty the bucket.')
LowRU: *Si gadis felimema si kemei.* ('The woman will iron the jeans.')
HighRU: *Si gadis tambamema si worel.* ('The woman will take the broccoli.')

Table 3.4.: Conditions Exp. 3
(Column 3: number of competitors; Column 4: potential learning mechanism(s))

Condition	Verb	Plausible Referents	Mechanism(s)
NoRU	restr.	1	strong SLCL
LowRU	restr.	2	weak SLCL & CSWL
HighRU	non-r.	4	CSWL

involved in a taking event). Participants were told that sentences were of the causative SVO form 'someone VERBs something'. They were asked to understand the sentences, and the instructions mentioned that this required knowing the unknown words.

Vocabulary Test For the forced-choice vocabulary tests (Phases 3 and 5) there were six depictions presented on the screen: the target (e.g., 'tomato') and another instance of the target's category (e.g., 'broccoli'), two objects of one of the two other categories (e.g., 'shirt' and 'skirt'), and two characters. In addition to recording the mouse clicks to the chosen referent, we introduced a confidence rating to have another, more sensitive measurement because there were only two nouns to be learned per condition: Participants were encouraged to press a number between 1 (not confident at all) and 9 (very confident) on the keyboard in order to indicate how sure they were about their choice of a referent.

3.3.3. Predictions

Firstly, we hypothesized that SLCL would be applied whenever it was available (even when it was not perfectly disambiguating referential uncertainty), facilitating noun learning. We moreover hypothesized that when SLCL was sufficient for learning (in Condition NoRU), it would be the only mechanism used. Both hypotheses were grounded on the way learners used verbal constraints on-line in Experiment 1: They unambiguously identified the target object during the verb and NP2, while ignoring all other objects. This predicted better learning rates and confidence ratings for Conditions NoRU and LowRU (where SLCL could be used) than for Condition HighRU (where only CSWL could be used). Concerning the comparison between learning rates and confidence ratings in NoRU and LowRU, we considered two hypotheses and the following predictions as plausible: Hypothesizing that, in LowRU, CSWL completely compensates for the low referential uncertainty left by SLCL while highly engaging the learner predicted that learning in LowRU would be as good as learning in NoRU. Hypothesizing that additionally applying CSWL would not completely compensate referential uncertainty left by SLCL but would require a higher cognitive effort, in contrast, predicted learning in LowRU to be worse than learning in NoRU.

Given that we expected a fast-mapping like situation in Condition NoRU which made considering distractor objects completely unnecessary, we predicted that eye-movements during NP2 (in Phases 2 and 4) would reflect a strong preference for inspecting the target. In Condition LowRU, in contrast, we expected verbal constraints to only narrow down the on-line search space from four to two. This predicts eye-movements to reflect both a preference for the target and a for the competitor. Finally, for Condition HighRU, we expected that the on-line search space contained four potential referents. Our prediction was therefore an equally strong consideration of all competitors, reflected by eye-movements.

Finally, we hypothesized that in Block 2 participants could exclude those objects as potential referents which have been already linked to a world-word-mapping in Block 1 (assuming the use of the principle of mutual exclusivity, Markman & Wachtel, 1988). This predicts an enhanced noun learning in Block 2 compared to Block 1.

3.3.4. Data Analysis & Results

3.3.4.1. Off-line Results

Noun learning (72%) was reliably better than chance (25%) across all conditions (see Table 3.6) and correlated positively with confidence ratings ($r_s = .430, p < .001$). Learning was significantly better in Block 2 than Block 1 (all conditions: $\chi(1) = 30.769, p < .001$; Condition NoRU: $\chi(1) = 6.309, p < .05$; Condition LowRU: $\chi(1) = 10.167, p < .01$; Condition HighRU: $\chi(1) = 16.568, p < .001$; see Table 3.5). Likewise, confidence ratings were higher in Block 2 than Block 1 (all conditions: $\chi(1) = 12.849, p < .001$; Condition NoRU: $\chi(1) = 5.416, p < .05$; Condition LowRU: $\chi(1) = 5.476, p < .05$; Condition HighRU: $\chi(1) = 10.688, p < .01$; see Table 3.6).

Table 3.5.: Noun learning percentages (t-tests against chance = 25%), Exp. 3

RU	Blocks 1+2	Block 1	Block 2
all	**72%**($t(62) = 12.18, p < .001$)	**62%**($t(62) = 6.90, p < .001$)	**83%**($t(62) = 14.24, p < .001$)
No	**77%**($t(62) = 10.04, p < .001$)	**69%**($t(62) = 6.24, p < .001$)	**85%**($t(62) = 12.56, p < .001$)
Low	**74%**($t(62) = 9.25, p < .001$)	**64%**($t(62) = 5.19, p < .001$)	**85%**($t(62) = 11.34, p < .001$)
High	**66%**($t(62) = 7.43, p < .001$)	**52%**($t(62) = 3.49, p < .001$)	**80%**($t(62) = 8.69, p < .001$)

Table 3.6.: Confidence ratings, Exp. 3

	Ref.Un.	Blocks 1+2	Block 1	Block 2
1	*all*	5.73	5.06	6.39
2	*NoRU*	6.98	6.34	7.50
3	*LowRU*	6.42	5.88	6.80
4	*HighRU*	5.40	4.45	6.02

We found the same tendencies in Blocks 1 and 2 (Figures 3.15, 3.16, and 3.14): Nouns were learned best and confidence was highest in Condition NoRU and worst in Condition HighRU. For inferential statistics, we analyzed both learning rates and confidence ratings with linear mixed-effects models, using logistic regression for the categorical learning rates (logit link function) and linear regression for the continuous confidence ratings, with participant and item as random factors (see Section 3.1.4.2 for explanations about this way of analyzing). For confidence ratings we additionally calculated Monte Carlo Markov Chain values (MCMCs) whose p-values are a good estimate of a factor's significance (but are only applicable for continuous variables, Baayen et al., 2008).

For learning rates, we found a main effect of factor Referential Uncertainty for Block 1 ($\chi(2) = 6.063, p < .05$) and a marginally significant difference between performance in Conditions NoRU and HighRU (see Table 3.7, Row 3) as well as a marginally significant difference between learning in Conditions LowRU and HighRU (see Table 3.7, Row 6). The difference between learning in Conditions NoRU and LowRU was not reliable (see Table 3.7, Row 2). There was no main effect of Referential Uncertainty in Block 2 ($\chi(2) = 0.831, p = .660$) nor in both blocks taken together ($\chi(2) = 4.289, p = .117$).

Importantly, we found main effects of Referential Uncertainty for confidence ratings for all parts: Block 1 ($\chi(2) = 15.647, p < .001$), Block 2 ($\chi(2) = 23.169, p < .001$), as well as

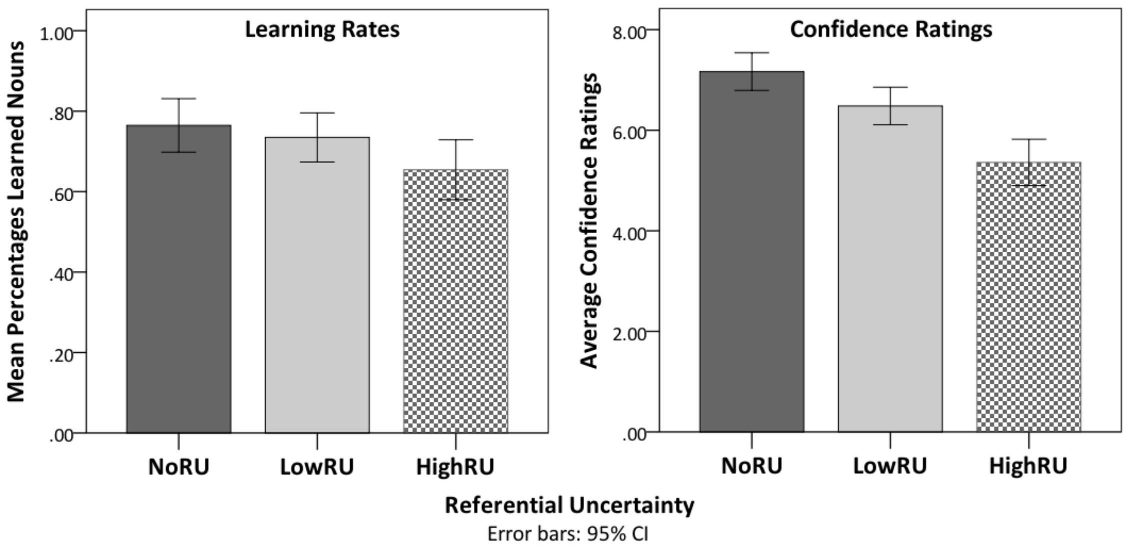


Figure 3.14.: Off-line data, Exp. 3, Blocks 1-2: Mean numbers of learned nouns (left chart) and average confidence ratings (right chart)

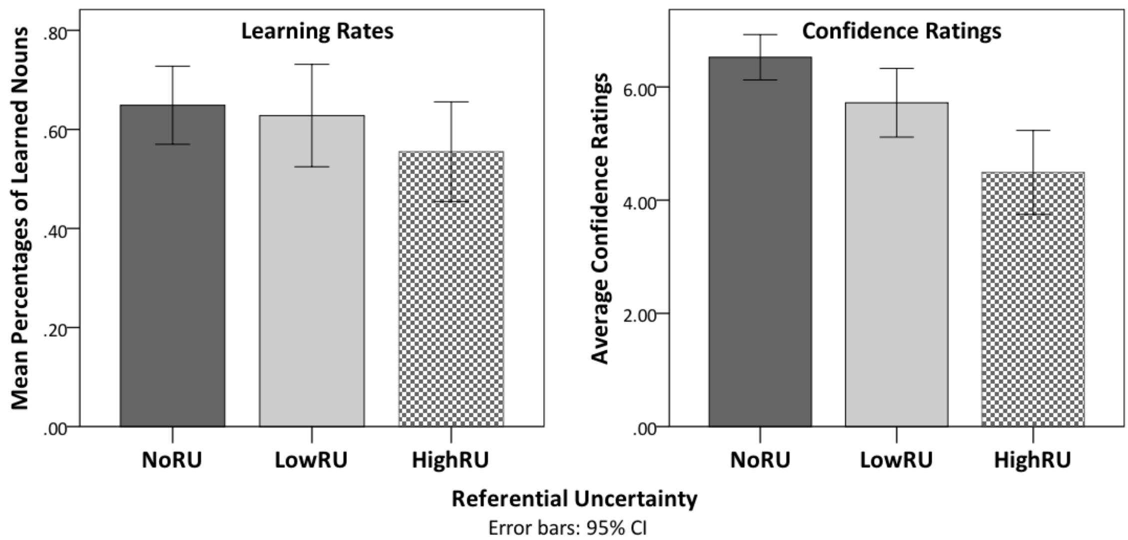


Figure 3.15.: Off-line data, Exp. 3, Block 1: Mean numbers of learned nouns (left chart) and average confidence ratings (right chart)

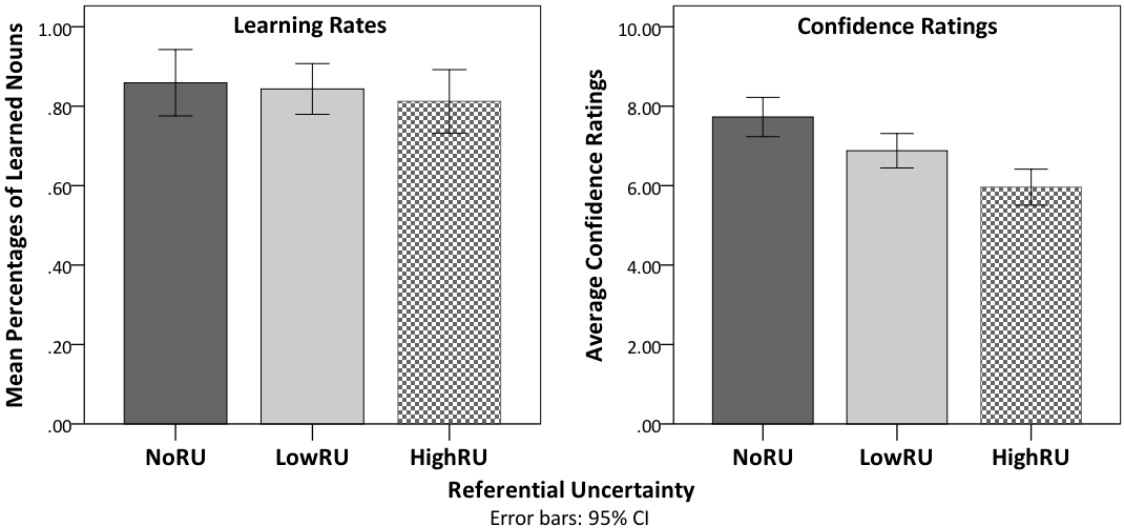


Figure 3.16.: Off-line data, Exp. 3, Block 2: Mean numbers of learned nouns (left chart) and average confidence ratings (right chart)

Table 3.7.: Lmer models for learning rates Exp. 3, Block 1
$learningrates \sim ReferentialUncertainty + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
1	(Int) (NoRU)	1.171	0.427	2.744	$< .010$
2	LowRU	-0.304	0.432	-0.704	$= .481$
3	HighRU	-1.065	0.429	-2.482	$= .050$
4	(Int) (LowRU)	0.867	0.418	2.075	$< .050$
5	NoRU	0.304	0.432	0.704	$= .481$
6	HighRU	-0.761	0.420	-1.813	$= .070$

Table 3.8.: Lmer models & p-values from MCMC sampling for confidence ratings, Exp 3, Blocks 1-2
$confidencerating \sim ReferentialUncertainty + (1|sub) + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
1	(Int) (NoRU)	7.063	0.379	18.648	7.041	.000	$< .001$
2	LowRU	−0.649	0.285	−2.274	−0.649	.038	$< .050$
3	HighRU	−1.758	0.301	−5.842	−1.706	.000	$< .001$
4	(Int) (LowRU)	6.415	0.381	16.831	6.416	.000	$< .001$
5	NoRU	0.649	0.285	2.274	.0625	.039	$< .050$
6	HighRU	−1.109	0.299	−3.701	−1.084	.001	$< .001$

Table 3.9.: Lmer models & p-values from MCMC sampling for confidence ratings, Exp 3, Block 1
$confidencerating \sim ReferentialUncertainty + (1|sub) + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
1	(Intercept) (NoRU)	6.495	0.456	14.243	6.393	.000	$< .001$
2	LowRU	−0.693	0.449	−1.545	−0.552	.284	$= .125$
3	HighRU	−2.087	0.490	−4.256	−1.939	.001	$< .001$
4	(Intercept) (LowRU)	5.802	0.467	12.418	5.835	.000	$< .001$
5	NoRU	0.693	0.449	1.545	.559	.0283	$= .125$
6	HighRU	−1.394	0.490	−2.848	−1.382	.012	$< .050$

Table 3.10.: Lmer models & p-values from MCMC sampling for confidence ratings, Exp 3, Block 2
$confidencerating \sim ReferentialUncertainty + (1|sub) + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
1	(Int.) (NoRU)	7.423	0.423	17.546	7.450	.000	$< .001$
2	LowRU	−0.648	0.322	−2.016	−0.642	.094	$< .050$
3	HighRU	−1.658	0.332	−4.990	−1.610	.000	$< .001$
4	(Int.) (LowRU)	6.774	0.424	15.968	6.813	.000	$< .001$
5	NoRU	0.648	0.322	2.016	0.634	.092	$< .050$
6	HighRU	−1.010	0.330	−3.061	−0.967	.014	$< .050$

both blocks together ($\chi(2) = 31.833, p < .001$). There were also interesting effects when comparing single conditions: When taking both blocks together (Table 3.8), there were significant differences between all conditions: NoRU and LowRU (Row 1), NoRU and HighRU (Row 3), as well as LowRU and HighRU (Row 6). The same pattern was found for Block 2 (Table 3.10): Significant comparisons between Conditions NoRU and LowRU (Row 2), between Conditions NoRU and HighRU (Row 3), and LowRU and HighRU (Row 6). In Block 1 (Table 3.9), confidence was significantly lower in Condition HighRU than in Condition NoRU (Row 3) and Condition LowRU (Row 6). That means that while we found the expected pattern in both learning rates and confidence ratings (best learning in NoRU and worst learning HighRU), effects in confidence ratings were much more reliable.

3.3.4.2. On-line Results

Eye-movements were analyzed as in Experiments 1 and 2 (see Section 3.1.3.2) except that the factor List was included as random factor. This was done because lists were not perfectly counterbalanced with regard to the condition-category relation, due to keeping the number of lists feasible: Nouns from each of the three object categories (food, clothing, containers) were presented in only two of three conditions across lists (e.g., food nouns belonged to Condition NoRU in four lists and to Condition LowRU in four lists but never to Condition HighRU). Firstly, we analyzed the looking pattern averaged across all conditions and both blocks to see whether participants generally followed the sentence structure and learned character and object names over time. Secondly, inspections during NP2 for single conditions were analyzed since we were mainly interested in differences between the resolution of the postverbal noun in the three conditions, and for single blocks to see learning development. Thirdly, eye movements during the verb region for single conditions were analyzed in order to examine whether there were anticipatory looks towards referents (averaged across blocks).

Considering all conditions and both blocks, the eye-movement pattern for all parts of the experiment resembles that of Experiment 1: We found reliably more inspections to the character than to all objects during NP1 (main effect ROI: $\chi(4) = 73.894, p < .001$; interaction of factors ROI and Referential Uncertainty:$\chi(8) = 19.438, p < .05$; Table 3.11, Rows 2-5). In the verb region, there were significantly more looks to the target object than to the other regions (main effect ROI: $\chi(4) = 50.207, p < .001$; interaction of factors Referential Uncertainty and ROI: $\chi(10) = 21.247, p < .05$; Table 3.11, Rows 7-10). There were also reliably more inspections on the target object than to the other regions in NP2 (main effect ROI: $\chi(4) = 280.910, p < .001$; interaction of ROI with Referential Uncertainty: $\chi(10) = 50.574, p < .001$; Table 3.11, Rows 12-15). See timegraph in Figure 3.17.

There were clear differences between conditions for inspections in NP2 : In Condition

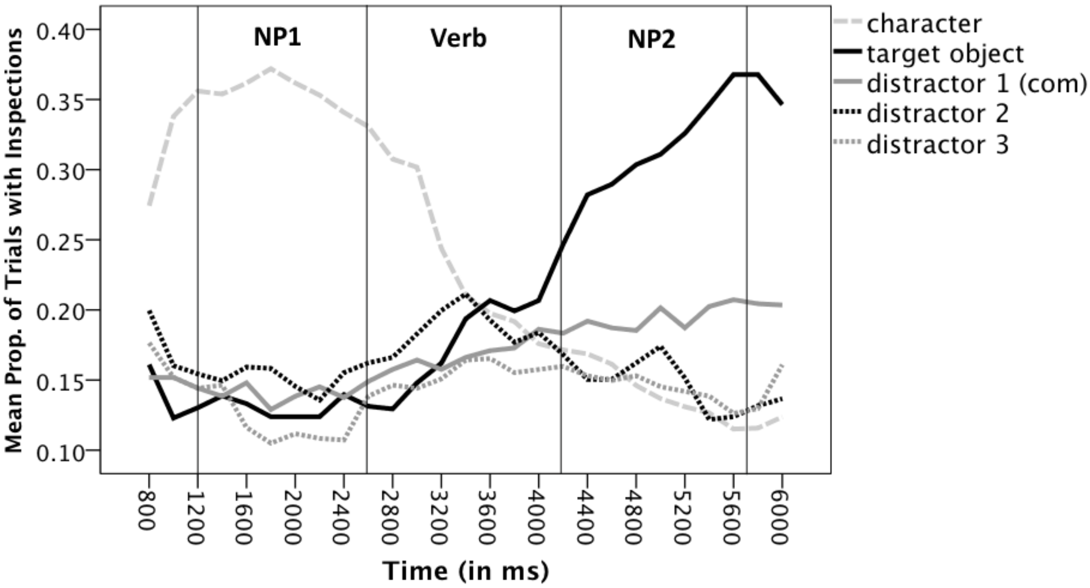


Figure 3.17.: Timegraph, Exp. 3

Table 3.11.: Lmer models for inspections on target vs. distractors, averaged across conditions and blocks, Exp. 3
$Inspections \sim ROI + (1|sub) + (1|item) + (1|list), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
NP1					
1	(Int) (char)	-0.401	0.114	-3.516	$< .001$
2	tar	-0.604	0.093	-6.488	$< .001$
3	di1	-0.640	0.094	-6.808	$< .001$
4	di2	-0.437	0.094	-4.649	$< .001$
5	di3	-0.690	0.100	-6.920	$< .010$
verb					
6	(Int) (tar)	-0.092	0.116	-0.789	$= .430$
7	char	-0.621	0.089	-7.000	$< .001$
8	di1	-0.297	0.091	-3.242	$< .010$
9	di2	-0.263	0.093	-2.832	$< .010$
10	di3	-0.348	0.096	-3.631	$< .001$
NP2					
11	(Int) (tar)	0.031	0.083	0.377	$= .706$
12	char	-1.496	0.098	-15.329	$< .001$
13	di1	-0.469	0.092	-5.107	$< .001$
14	di2	-0.942	0.097	-9.665	$< .001$
15	di3	-0.933	0.100	-9.331	$< .001$

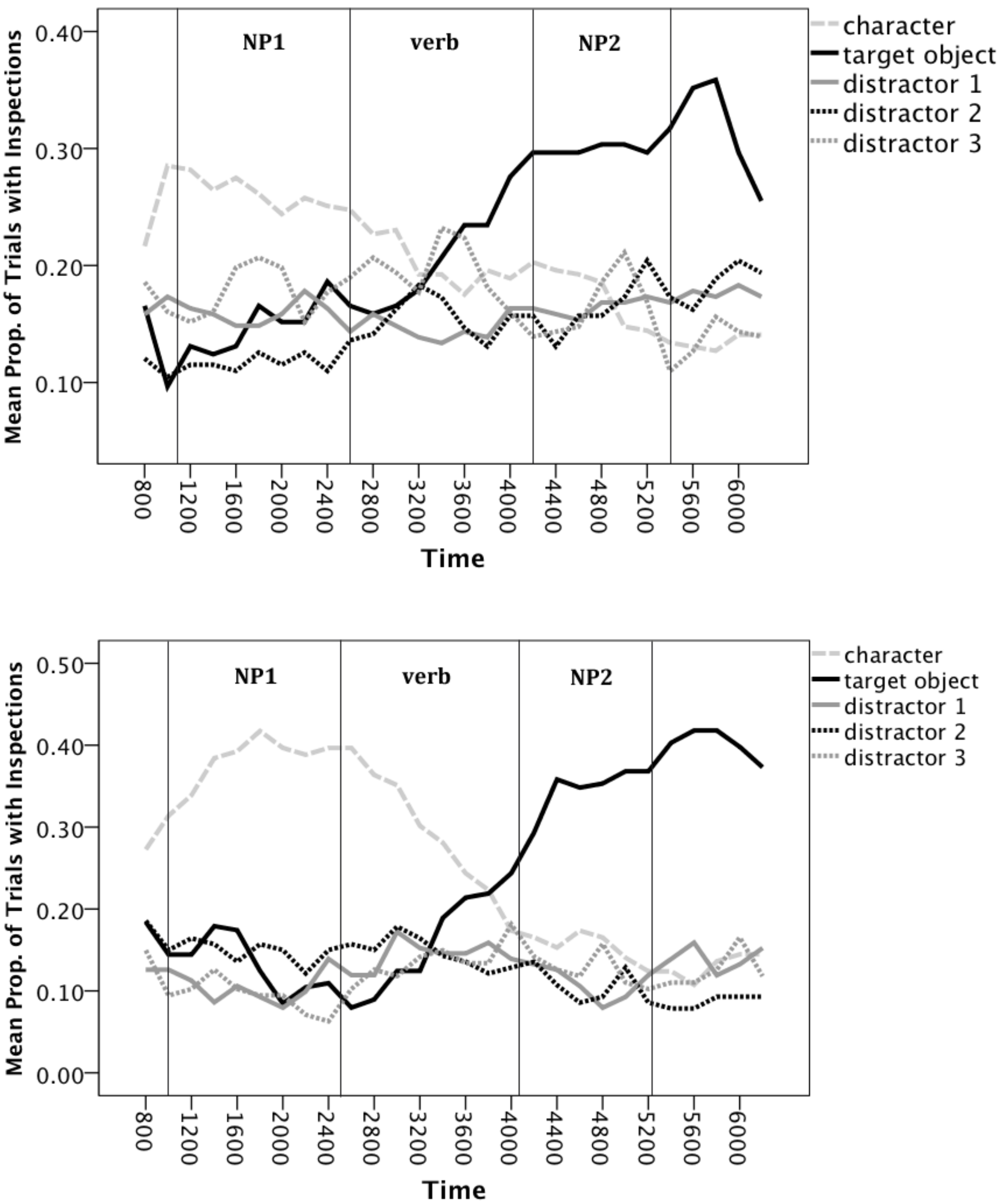


Figure 3.18.: Timegraph, Exp. 3, Block 1 (top) and Block 2 (bottom), Condition NoRU

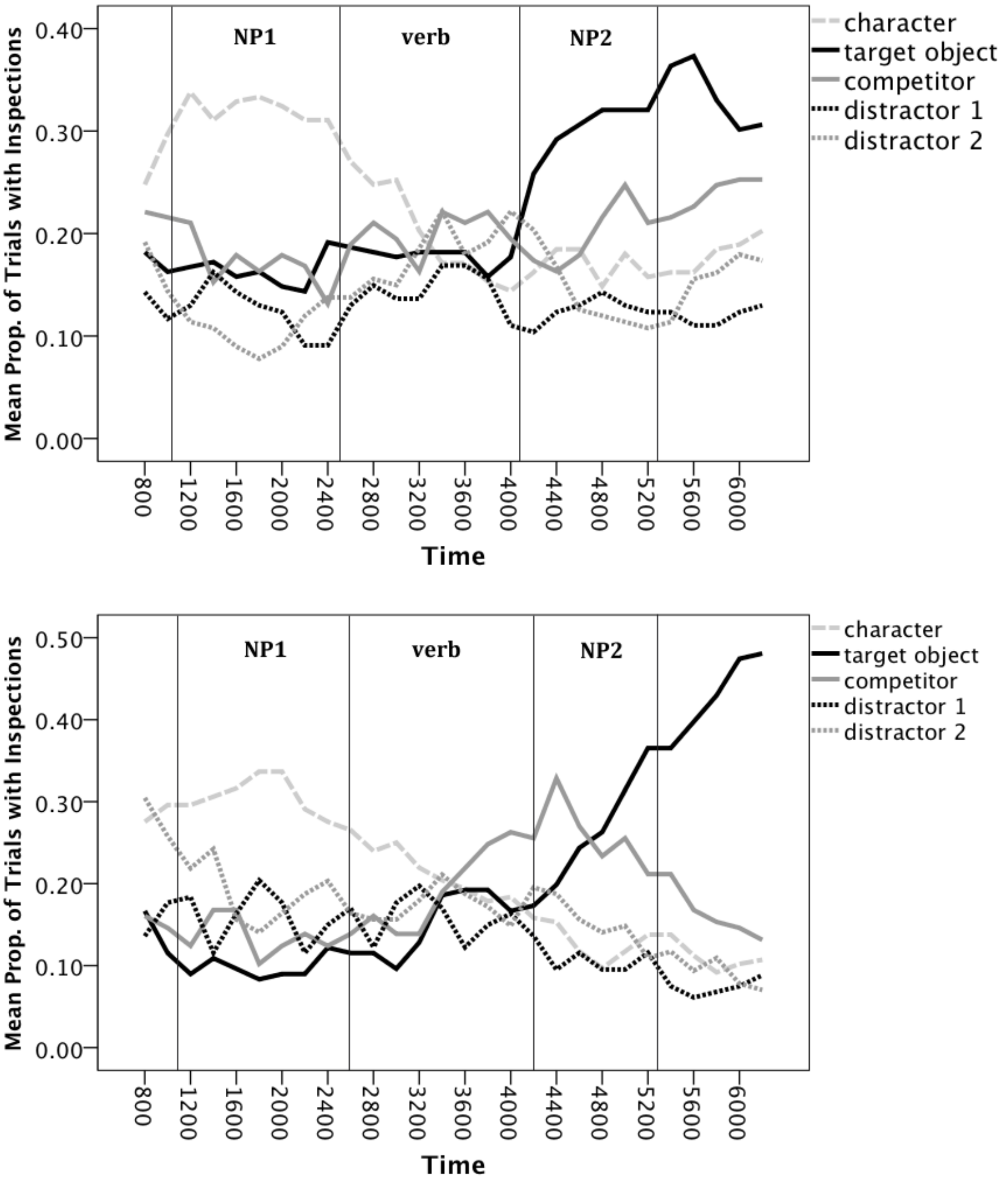


Figure 3.19.: Timegraph, Exp. 3, Block 1 (top) and Block 2 (bottom), Condition LowRU

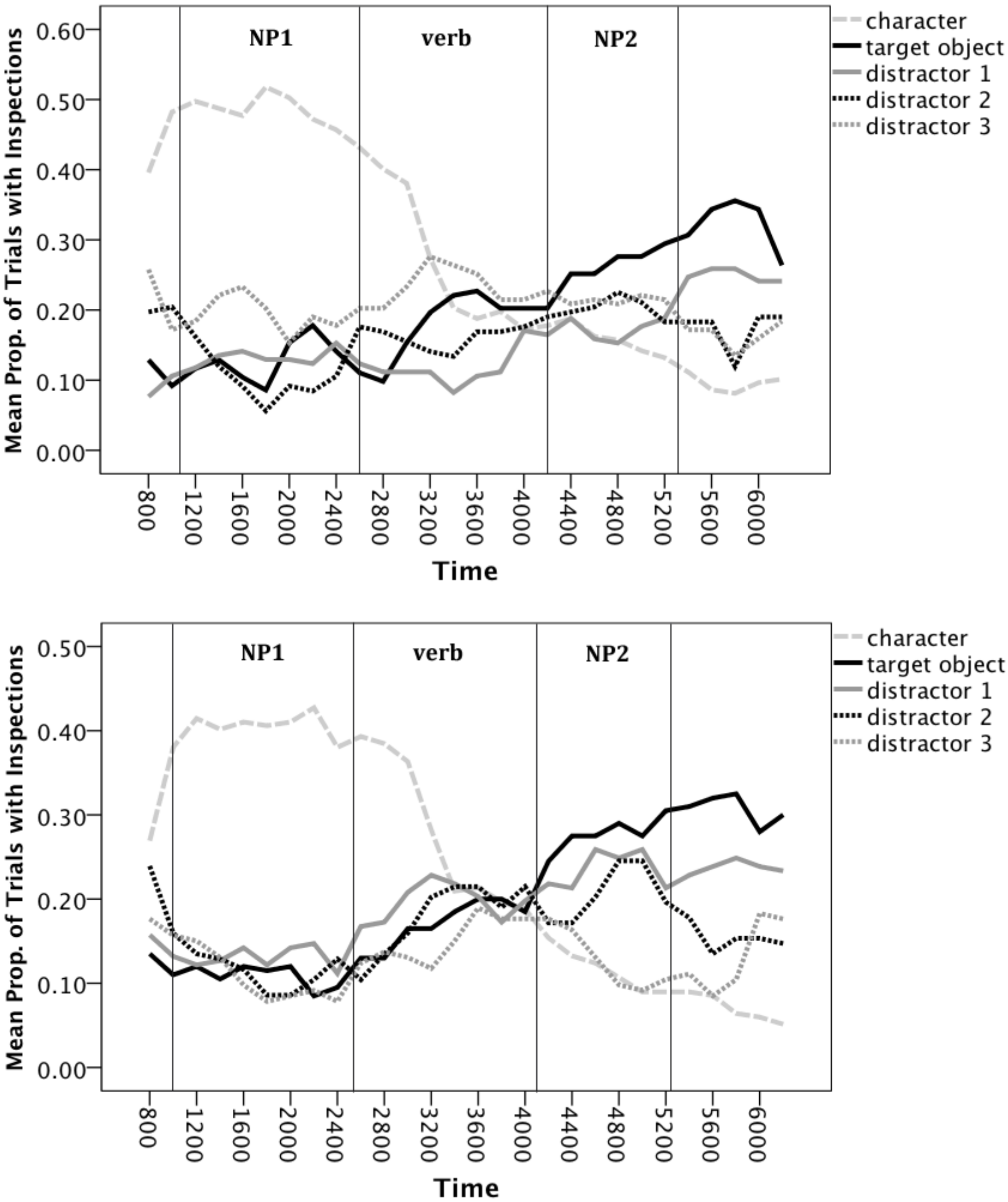


Figure 3.20.: Timegraph, Exp. 3, Block 1 (top) and Block 2 (bottom), Condition HighRU

NoRU, the target object was inspected reliably more often than the character and the distractor objects in Block 2 (main effect ROI: $\chi(4) = 52.280, p < .001$; Table 3.14, Rows 2-5) and both blocks together (main effect ROI: $\chi(4) = 69.236, p < .001$; Table 3.12, lines 2-5). In Block 1 the difference to one (random) distractor was not significant (main effect ROI: $\chi(4) = 30.376, p < .001$; Table 3.13, Rows 2-5). See barchart in Figure 3.21 timegraphs in Figure 3.18.

In Condition LowRU, the target was inspected significantly more often than the other objects in Block 1 (main effect ROI: $\chi(4) = 50.400, p < .001$; Table 3.13, Rows 12-15), Block 2 (main effect ROI: $\chi(4) = 73.879, p < .001$; Table 3.14, Rows 12-15), and both blocks (main effect ROI: $\chi(4) = 118.750, p < .001$; Table 3.12, Rows 12-15), except that the difference between target and competitor in Block 1 was not significant (Table 3.13, Row 13). Moreover, the competitor was looked at reliably more than the other objects (except the target) in Block 1 (Table 3.13, Rows 17-20), Block 2 (Table 3.14, Rows 17-20), and both blocks (Table 3.12, Rows 17-20). See barchart in Figure 3.21 and timegraphs in Figure 3.19.

For Condition HighRU, the target was inspected reliably more often than the character and the distractors in Block 1 (main effect ROI: $\chi(4) = 66.374, p < .001$; Table 3.13, Rows 22-25), Block 2 (main effect ROI: $\chi(4) = 76.778, p < .001$; Table 3.14, Rows 22-25), and both blocks (main effect ROI: $\chi(4) = 138.650, p < .001$; Table 3.12, Rows 22-25), except that the difference between looks to the target and to one of the three distractors was non-significant in all parts (Tables 3.12 - Table 3.14, Rows 17-20). See barchart in Figure 3.21 timegraphs in Figure 3.20. We found posthoc that this distractor was exactly that one which always shared the semantic category with the target (e.g., when the target was bucket, the distractor was another container). There were also significantly more looks to this distractor than to the other distractors and the character in both blocks (Table 3.12, Rows 27-30).

We additionally compared eye-movements during NP2 in Block 1 and Block 2 (separate for the three conditions) in order to examine whether identification of referents improved over time. We found that the target object was inspected more often in Block 2 than Block 1, as reflected in proportions. This was true for all conditions, with significance reached in Condition NoRU ($\chi(1) = 15.903, p < .001$; 3.15, Row 2) and Condition HighRU ($\chi(1) = 5.911, p < .05$; Row 6) and marginal significance approached in Condition LowRU ($\chi(1) = 2.283, p = .131$, Row 4). The data revealed, moreover, that there was a marginal interaction between factors Block and Referential Uncertainty for proportions of looks to the target object during NP2 ($\chi(2) = 5.594, p = .061$): The largest increase in looks to the target from Block 1 to Block 2 happened in Condition NoRU.

Given that we were not interested in potential developmental changes of eye movements

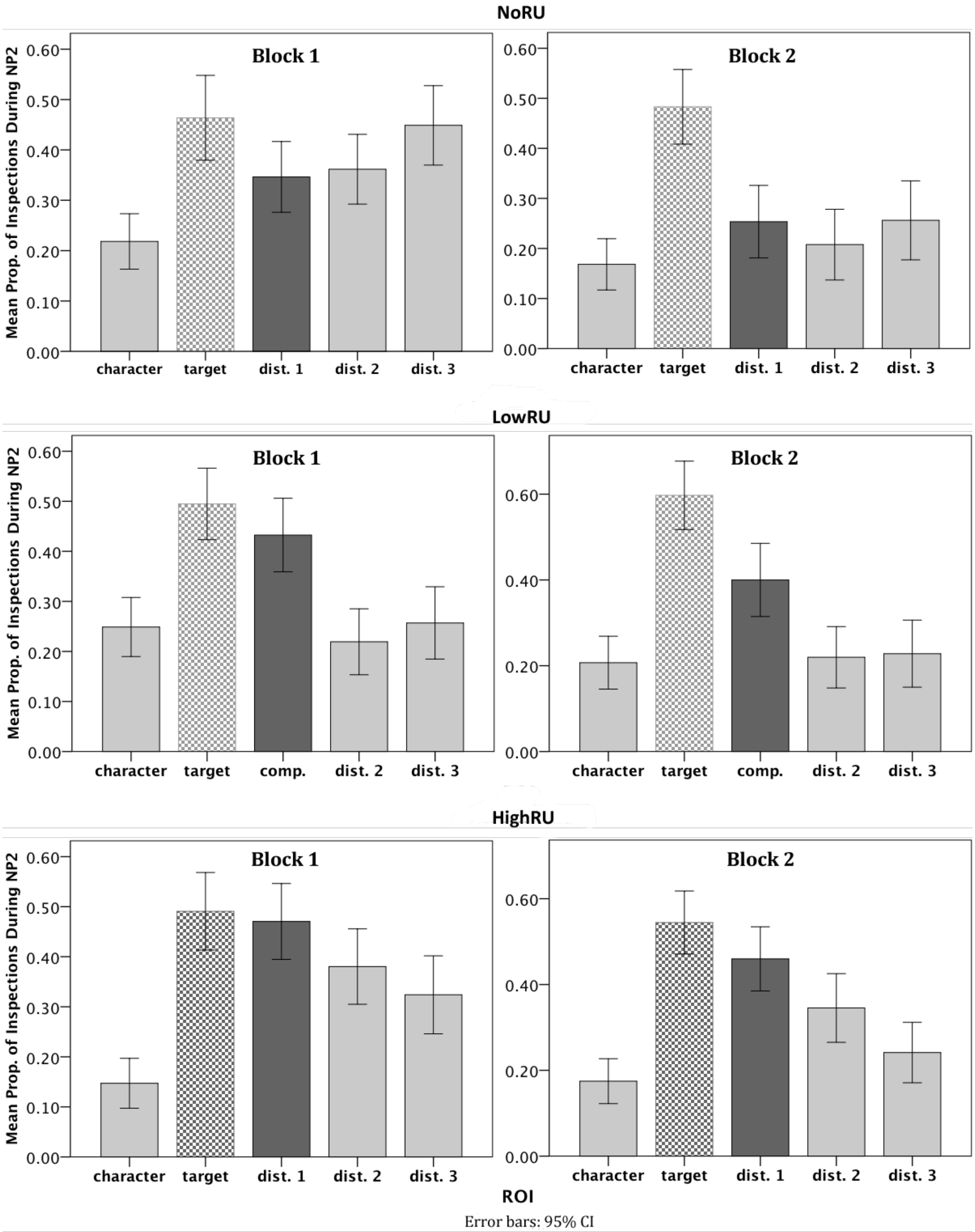


Figure 3.21.: Proportions of trials with at least one inspection during NP2, Exp. 3, Condition NoRU (top), Condition LowRU (mid), and Condition HighRU (bottom): Block 1 (left charts) and Block 2 (right charts)

Table 3.12.: Lmer models for inspections on target vs. distractors and distractor1/competitor vs. other ROIs during NP2, Exp. 3, Blocks 1-2

$InspectionsduringNP2 \sim ROI + (1|sub) + (1|item) + (1|list), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
NoRU					
1	(Int) (tar)	−0.116	0.136	−0.845	= .393
2	char	−1.342	0.169	−7.959	< .001
3	di1	−0.698	0.167	−4.172	< .001
4	di2	−0.746	0.168	−4.430	< .001
5	di3	−0.460	0.170	−2.707	< .010
6	(Int) (di1)	−0.814	0.143	−5.711	< .001
7	tar	0.698	0.167	4.172	< .001
8	char	−0.644	0.174	−3.702	< .001
9	di2	−0.047	0.174	−0.273	= .785
10	di3	0.238	0.175	1.359	= .174
LowRU					
11	(Int) (tar)	0.157	0.121	1.294	= .196
12	char	−1.381	0.164	−8.407	< .001
13	com	−0.505	0.159	−3.169	< .010
14	di2	−1.449	0.180	−8.046	< .001
15	di3	−1.308	0.182	−7.194	< .001
16	(Int) (com)	−0.348	0.128	−2.711	< .010
17	tar	0.505	0.159	3.169	< .010
18	char	−0.876	0.169	−5.181	< .001
19	di2	−0.944	0.184	−5.128	< .001
20	di3	−0.803	0.186	−4.316	< .001
HighRU					
21	(Int) (tar)	0.048	0.135	0.352	= .725
22	char	−1.743	0.175	−9.963	< .001
23	di1	−0.208	0.154	−1.350	= .177
24	di2	−0.660	0.163	−4.063	< .001
25	di3	−1.046	0.171	−6.102	< .001
26	(Int) (di1)	−0.161	0.135	−1.189	= .234
27	tar	0.208	0.1542	1.350	= .177
28	char	−1.535	0.175	−8.775	< .001
29	di2	−0.452	0.163	−2.778	< .010
30	di3	−0.838	0.172	−4.866	< .001

Table 3.13.: Lmer models for inspections on target vs. distractors and distractor1/competitor vs. other ROIs during NP2, Exp. 3, Block 1

$InspectionsduringNP2 \sim ROI + (1|sub) + (1|item) + (1|list), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
NoRU					
1	(Int) (tar)	−0.178	0.179	−0.996	= .319
2	char	−1.123	0.238	−4.714	< .001
3	di1	−0.479	0.234	−2.049	< .050
4	di2	−0.421	0.230	−1.831	= .067
5	di3	−0.060	0.236	−0.256	= .798
6	(Int) (di1)	−0.657	0.164	−4.000	< .001
7	tar	−0.479	0.234	2.049	< .050
8	char	−0.645	0.228	−2.830	< .010
9	di2	−0.057	0.220	0.262	= .794
10	di3	0.418	0.226	1.852	= .064
LowRU					
11	(Int) (tar)	−0.018	0.160	−0.113	= .910
12	char	−1.114	0.217	−5.139	< .001
13	com	−0.259	0.210	−1.232	= .218
14	di2	−1.266	0.242	−5.202	< .001
15	di3	−1.070	0.241	−4.449	< .001
16	(Int) (com)	−0.277	0.167	−1.661	= .100
17	tar	−0.259	0.210	1.232	= .218
18	char	−0.855	0.222	−3.848	< .001
19	di2	−1.007	0.247	−4.073	< .001
20	di3	−0.811	0.246	−3.304	< .001
HighRU					
21	(Int) (tar)	−0.050	0.172	−0.291	= .771
22	char	−1.753	0.257	−6.810	< .001
23	di1	−0.098	0.221	−0.444	= .657
24	di2	−0.498	0.227	−2.194	< .050
25	di3	−0.722	0.240	−3.013	< .010
26	(Int) (di1)	−0.419	0.169	−0.878	= .380
27	tar	0.098	0.221	0.444	= .657
28	char	−1.655	0.255	−6.497	< .001
29	di2	−0.400	0.224	−1.786	= .074
30	di3	−0.624	0.239	−2.615	< .010

Table 3.14.: Lmer models for inspections on target vs. distractors and distractor1/competitor vs. other ROIs during NP2, Exp. 3, Block 2
$InspectionsduringNP2 \sim ROI + (1|sub) + (1|item) + (1|list), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
NoRU					
1	(Int) (tar)	−0.048	0.166	−0.292	= .770
2	char	−1.571	0.240	−6.522	< .001
3	di1	−1.045	0.246	−4.243	< .001
4	di2	−1.301	0.265	−4.901	< .001
5	di3	−1.024	0.259	−3.958	< .001
16	(Int) (di1)	−1.093	0.206	−5.319	< .001
17	tar	−1.045	0.246	4.243	< .001
18	char	−0.526	0.270	−1.951	= .051
19	di2	−0.256	0.292	−0.3879	= .380
20	di3	0.021	0.286	0.074	= .941
LowRU					
6	(Int) (tar)	0.383	0.185	2.078	< .038
7	char	−1.774	0.256	−6.924	< .001
8	com	−0.829	0.247	−3.356	< .001
9	di2	−1.717	0.272	−6.304	< .001
10	di3	−1.655	0.282	−5.875	< .001
21	(Int) (com)	−0.446	0.200	−2.266	< .050
22	tar	0.829	0.247	3.356	< .001
23	char	−0.945	0.264	−3.574	< .001
24	di2	−0.888	0.279	−3.181	< .010
25	di3	−0.826	0.289	−2.858	< .010
HighRU					
11	(Int) (tar)	0.168	0.160	1.049	= .029
12	char	−1.753	0.238	−7.355	< .001
13	di1	−0.343	0.214	−1.602	= .109
14	di2	−0.839	0.234	−3.583	< .001
15	di3	−1.341	0.246	−5.443	< .001
26	(Int) (di1)	−0.175	0.163	−1.074	= .283
27	tar	0.343	0.214	1.602	= .109
28	char	−1.410	0.240	−5.878	< .001
29	di2	−0.495	0.236	−2.101	< .050
30	di3	−1.000	0.248	−4.025	< .001

Table 3.15.: Lmer models & p-values from MCMC sampling for proportions of inspections to target object during NP2 in blocks, Exp. 3

$inspections \sim block + (1|sub) + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
NoRU							
1	(Int.) (Block1)	6.461	0.447	14.442	7.231	$< .001$	$< .001$
2	Block2	0.0865	0.368	2.352	1.919	$< .050$	$< .050$
LowRU							
3	(Int.) (Block1)	6.090	0.464	13.124	6.816	$< .001$	0.0000
4	Block2	0.773	0.329	2.353	1.718	$< .050$	$< .050$
HighRU							
5	(Int.) (Block1)	4.485	0.487	9.206	5.434	$< .001$	$< .001$
6	Block2	1.555	0.459	3.388	2.603	$< .010$	0.010

during verb region over the course of the experiment, these inspections were analyzed averaged across both blocks only (see Figure 3.22). We found a main effect of ROI in all conditions (Condition NoRU: $\chi(4) = 20.374, p < .001$, Condition LowRU: $\chi(4) = 15.082, p < .010$, Condition HighRU: $\chi(4) = 36.544, p < .001$). In Condition NoRU, the target object was looked at reliably more often than the rest (Table 3.16, Rows 2-5). In Condition LowRU, the target object was inspected significantly more often than the character (but not the distractors, Table 3.16, Rows 12-15) but the competitor was also inspected significantly more often than the character (and the distractors, Table 3.16, Rows 17-20). In Condition HighRU, the target object was looked at reliably more often than the other regions except the distractor of the target's category (Table 3.16, Rows 12-25). The distractor which shared the semantic category with the target was looked at reliably more than the character (Table 3.16, Row 28).

3.3.5. Summary and Discussion

To summarize, off-line results of Experiment 3 reveal a tendency for learning rates and confidence ratings to be highest in Condition NoRU and lowest in Condition HighRU. Marginally significant differences in learning rates were found in Block 1 between Conditions NoRu and HighRU and between Conditions LowRU and HighRU. Confidence ratings in Block 1 and Block 2 reveal significant differences between Conditions NoRU and HighRU and between Conditions LowRu and HighRU; the difference between Conditions NoRU

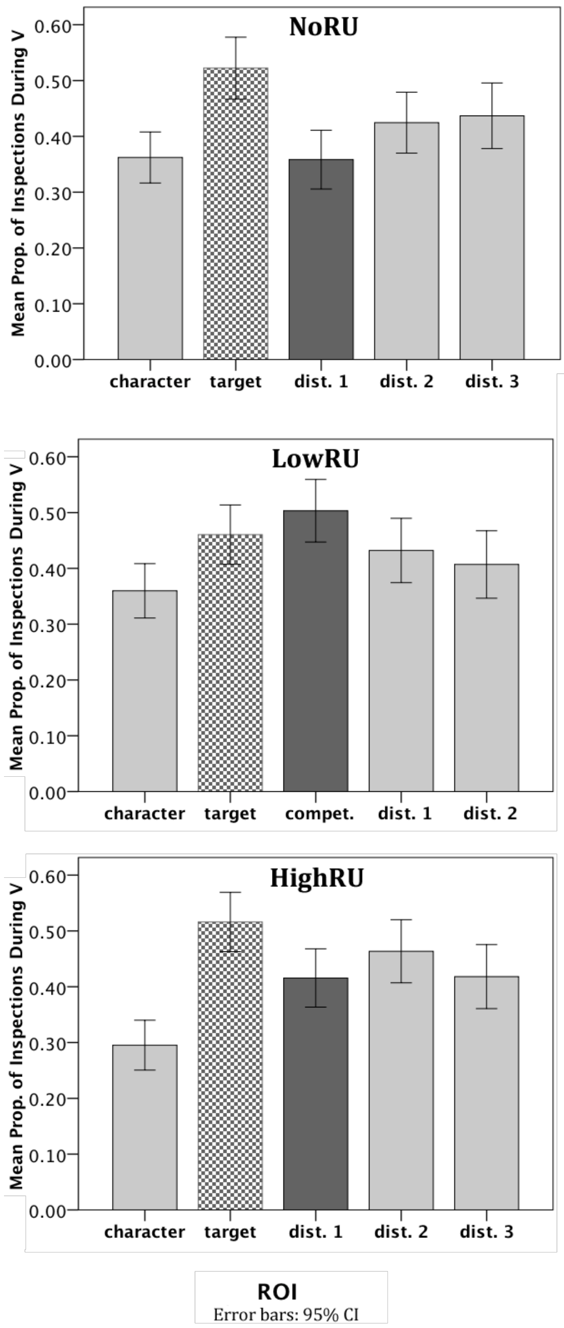


Figure 3.22.: Proportions of trials with at least one inspection to ROIs during the verb (averaged across blocks): Condition NoRU (top chart), Condition LowRU (mid chart), and Condition HighRu (bottom chart)

Table 3.16.: Lmer models for inspections on target vs. distractors and distractor1/competitor vs. other ROIs, during the verb, Exp. 3, Blocks 1-2

$$Inspections \sim ROI + (1|sub) + (1|item) + (1|list), family = binomial(link = "logit")$$

	Predictor	Coef.	SE	Wald z	p
NoRU					
1	(Int) (tar)	0.016	0.158	0.104	= .917
2	char	−0.604	0.153	−3.937	< .001
3	di1	−0.634	0.164	−3.874	< .001
4	di2	−0.395	0.161	−2.450	< .050
5	di3	−0.337	0.167	−2.024	< .050
6	(Int) (di1)	−0.618	0.159	−3.883	< .001
7	tar	0.634	0.164	3.874	< .001
8	char	0.030	0.156	0.191	= .848
9	di2	0.239	0.164	1.458	= .145
10	di3	0.296	0.169	1.750	= .080
LowRU					
11	(Int) (tar)	−0.220	0.153	−1.439	= .151
12	char	−0.412	0.154	−2.670	< .010
13	com	0.142	0.160	0.889	= .374
14	di2	−0.161	0.163	−0.988	= .323
15	di3	−0.274	0.169	−1.642	= .104
16	(Int) (com)	−0.078	0.158	−0.495	= .621
17	tar	−0.142	0.160	−0.889	= .374
18	char	−0.554	0.158	−3.496	< .001
19	di2	−0.303	0.167	−1.820	= .069
20	di3	−0.416	0.172	−2.416	< .050
HighRU					
21	(Int) (tar)	−0.021	0.147	−0.145	= .885
22	char	−0.897	0.155	−5.789	< .001
23	di1	−0.397	0.155	−2.556	< .050
24	di2	−0.237	0.159	−1.485	= .138
25	di3	−0.420	0.163	−2.580	< .010
26	(Int) (di1)	−0.419	0.148	−2.820	< .010
27	tar	0.397	0.155	2.556	< .050
28	char	−0.500	0.156	−3.206	< .010
29	di2	0.160	0.161	0.994	= .320
30	di3	−0.023	0.165	−0.142	= .887

and LowRU was significant in Block 2 and both blocks together. Moreover, learning rates and confidence ratings in all conditions are significantly higher in Block 2 than in Block 1. Eye-movements reveal that, collapsed over conditions, learners looked more often at the character than at the objects during NP1, while the target object was inspected significantly more often than the other ROIs during the verb and NP2. During the verb, we found a clear preference to inspect the target object for NoRU and tendencies to look at both the target object and the competitor object more often than at the other ROIs in Conditions LowRU and HighRU. During NP2, we found different patterns for the three conditions: In NoRU, we found a clear preference for inspections to the target object; differences between looks to the target and looks to all other ROIs were (marginally) significant in both blocks, except that one (random) distractor was not looked at significantly more often than the target in Block 1. In LowRU, in contrast, learners tended to look more often at both the target object and the competitor object. In Block 1, both target and competitor were inspected significantly more often than the other ROIs but the difference between target and competitor was not significant. The same effects were found for Block 2, except that the target was now looked at significantly more frequently than the competitor. In HighRU, the target object and the distractor object which belonged to the same semantic category as the target were looked at most frequently. While both were looked at significantly more often than the other ROIs in both blocks, the difference between the two was significant in neither block. Finally, we found that the target object was looked at proportionally more often in Block 1 than Block 2 during NP2; this was true for all conditions with the most pronounced difference between blocks in NoRU.

As expected, learning in Conditions NoRU and LowRU was better than learning in Condition HighRU as reflected by learning rates in Block 1 and confidence ratings in both blocks. These differences reveal the predicted learning boost caused by SLCL (verbal restrictions). Results concerning the comparison between learning in NoRU and LowRU are not entirely conclusive: Confidence ratings between Conditions NoRU and LowRU differ significantly only in Block 2 and no effect was found for learning rates. The fragile difference may indicate that while applying CSWL in Condition LowRU compensated part of the referential uncertainty that remained based on SLCL, compensation was not perfect. This interpretation would support a hypothesis which lies between the two hypotheses that we defined. However, given that there were only two nouns in each condition per block, it is also possible that missing significances are not meaningful but subject to data sparsity. Either way, it is clear and insightful that, in Condition LowRU, SLCL and CSLW were used in a complementary way and that CSWL at least partially compensated for the referential uncertainty that SLCL left.

The missing significances in learning rates in Block 2 (and due to this also in the collapsed data of both blocks) are not very surprising and most likely due to a ceiling effect (learning in all conditions was high). This effect was probably caused by learners' applying of the principle of mutual exclusivity: They used their knowledge from Block 1 to exclude objects as referents. Data sparsity may also have contributed to this ceiling effect. Importantly, however, the more sensitive confidence ratings reveal that all pairwise differences between conditions in Block 2 and both blocks together are significant.

The on-line data reveals that, averaged across all conditions and both blocks, eye-movements reflect following of the sentence structure: Learners made referential looks during NP1 and NP2 and anticipatory looks in the verb region. More interestingly, differences between conditions in (referential) inspections during NP2 mainly support the off-line results. Whereas only the target object was preferred in Condition NoRU, both objects belonging to the target object's semantic category showed a preference for being inspected in Condition LowRU. The finding that the difference between looks to the target object and looks to the competitor object in LowRU was not significant in Block 1 but were in Block 2 nicely supports that learning improved from Block 1 to Block 2. The only surprising finding in eye-movements during NP2 in NoRU-trials is that, in Block 1, the difference between looks to the target object and looks to one (random) distractor object is not significant. We attribute this results to noise, indicating that eye-movements in Experiment 3 should be interpreted carefully: It seems plausible that the cognitive requirements for simultaneously following a spoken sentence rapidly and precisely, on the one hand, and for conducting complex CSWL, on the other hand, are different. As both processes potentially affect eye-movements in different ways, they may, together, result in an unclear pattern. While processing a spoken sentence includes highly time-locked integration of visual information, CSWL per definition proceeds across situations and is based on associating, memorizing, and comparing visual information. Given that this is the first eye-tracking CSWL-study with nouns which are embedded in sentences, we cannot make any direct comparisons to previous results.

The gaze pattern for Condition HighRU (preferences for the target as well as for the distractor which shared its category) is unforseen but interesting: Although verbs in their originally trained sense did not constrain the category of the object referent, the distractor with the same category as the target was preferred over the other distractors (which were of categories associated with other restrictive verbs). The non-restrictive verb *tambamema* ('to take'), for instance, was used with objects of all three categories in the verb-learning phase. Moreover, the most obvious German equivalent to this verb, *nehmen*, is not restrictive in any way. However, in the noun-learning phase, *tambamema* always co-occurred with a noun which referred to a container (and therefore there were also always container objects in the scene when the verb *tambamema* was used). We therefore consider eye-movements as evidence for the theory that participants learned a new co-occurrence restriction for non-restrictive verbs during noun learning (e.g., container objects and 'take') simply by recognizing the co-occurrence of verbs and categories and, probably, by excluding the objects which were already selected by another verb. This behavior suggests an interesting statistical learning strategy, associating object categories and verbs.

Block analysis of eye-movements during NP2 reveals that referent identification was better in Block 2 than Block 1 in all conditions. This suggests that learning in Block 2 was boosted based on the knowledge that learners had already acquired in Block 1 while relying on the principle of mutual exclusivity. The result that the increase in learning from Block 1 to Block 2 was highest in Condition NoRU reflects that learning in NoRu was easiest, supporting the off-line results.

Data from eye-movements during the verb reveals a clear anticipatory preference to inspect the target object in Condition NoRU. That indicates that learners exploited verbal restrictions to identify the target even before it was named. In Conditions LowRU and HighRU, we found a tendency to inspect both target and competitor. For Condition LowRU, this pattern is fully in accordance with our predictions: While the verb constrained the semantic category of the object noun, there were two plausible referents in the scene. The anticipatory effect in HighRU-trials fit in with the above discussed observations made for looks in NP2 and show that the newly learned verbal restrictions were even used rapidly enough to anticipate referents.

Taken all results together, the third experiment clearly reveals that on-line identification and nouns learning was best when SLCL was available (in NoRU and LowRU). Trials in Condition NoRU further supported earlier findings (from Experiments 1 and 2) that SLCL can narrow down the number of potential referents to one, a situation close to fast-mapping. Learning in Condition LowRU moreover clearly reveals that both SLCL and CSWL were applied in a complementary way: CSWL compensated for the referential uncertainty that remained based on SLCL. The fragile difference in learning and confidence between NoRu and LowRU may be due to data sparsity. However, we do not exclude the alternative explanation that the difference may have been decreased by learners' extra engagement in LowRU, caused by using CSWL in addition to SLCL. Eye-movements during NP2 and the verb in all conditions confirm the off-line results except that there was an unexpected preference to look at both members of the target category in Condition HighRU. We attribute this to spontaneous learning of novel verbal restrictions. Finally, we found a boost in learning over time, manifested in differences between Block 1 and 2 effect which was best for Condition NoRU.

3.4. Interim Summary

Experiments 1 to 3 reveal that CSWL and SLCL (in particular, learning based on verbal constraints) are both useful and powerful mechanisms for adult language novices in non-instructed environments who want to understand and learn second language input: They assist language novices in on-line comprehension and enable them to gain the command on a set of 12-14 noun meanings within time periods as short as about 30 minutes. Additionally, they interact in that they complement each other. Results from experiments 1 to 3 further shed light on the on-line sentence comprehension of beginning adult language learners: Participants succeeded quickly in processing sentences incrementally, predictively, and multi-modally, very similar to the way native speakers do.

We found in Experiment 1, that adult language learners conducted CSWL and SLCL when novel nouns were embedded in sentences and referents were parts of visual scenes. Data also reveal that SLCL (in particular verbal restrictions), together with people's world knowledge and the visual scene, helped learners to identify referents for novel words on-line and that these verbal restrictions were exploited rapidly and even predictively in a native-like way. Finally, our results at least support the hypothesis that SLCL also boosts noun learning. These findings were replicated in Experiment 2 which further reveals that predicting referents based on verbal constraints may itself enhance noun learning. Results of Experiment 2 also provide ideas about learners' on-line processing of sentences with OVS word order. Experiment 3 provided evidence for the hypotheses that SLCL boosts noun learning and that SLCL interacts with CSWL when both provide complementary information regarding the learning of a single noun. Specifically, we found that CSWL can compensate for the referential uncertainty that remains after SLCL.

While these are interesting and novel findings, some questions concerning the interplay of CSWL and SLCL remain. Firstly, while we know that both can interact in jointly assigning a noun's meaning, it is unclear how both mechanisms are used when they support conflicting meanings. The interplay of CSWL and SLCL in a conflicting situation potentially also sheds light on their priority: If one of both mechanisms is stronger than the other, it should also have more influence on the learner's decision or even completely override the other mechanism. Secondly, it is important to further investigate whether both learning mechanisms are still applicable when learning scenarios become noisy in a more natural way, that is, when the visual scene does not reliably contain all referents and verb information can only be useful when being transferred to other situations. These issues are the focus of the experiments reported in the next chapter.

4. CSWL and SLCL in Conflict

The learning setting developed in Experiments 1 to 3 offers a more natural setting than other word learning experiments in that, firstly, words are embedded in linguistic contexts and referents are embedded in embedded in visual contexts, secondly, different learning cues and mechanisms are considered in parallel. However, it still suffers from another kind of idealization, namely that different cues are fully in accordance with one another and, further, that cues which need to be integrated together to help word learning (*co-cues*) are always available simultaneously (e.g., verb information and visual scene containing plausible referents). In realistic learning settings, this is not necessarily the case: Learning cues are imperfect, information is frequently ambiguous and sometimes conflicting. It is therefore important to examine how helpful different cues are when in conflict, how they influence each other's use, and which are prioritized over others. Additionally, investigating CSWL and SLCL when they are in conflict potentially sheds further light on their nature and interplay: If one of both is more dominant, it should override the other.

Imagine again you are a learner of English. You join a conversation between two girls in a park and one of them says 'Absolutely, my nephew loves playing with daxes.' while looking at some flowers. You understand the sentence's structure and all of the words contained in it except for the noun *daxes*. You do not now what the conversation was about, so the broader linguistic context does not help you. What the sentential context tells you, however, is that a *dax* must be something that people can play with. The visual context does potentially provide information as well: *daxes* could refer to the flowers (maybe it is the name of that particular type of flower). However, you might find that not very likely because children do not typically tend to play with flowers. But still, it is possible to, for instance, arrange flowers and someone could call that playing (in a language that you have no perfect command of you could also just not know all ways how the verb *play* can be used). Now imagine that two days later, you are in a gift shop and you hear a man asking the shop assistant 'I see. But how about the daxes?' pointing to one shelf which contains uncommonly shaped vases, artificial flowers, and something that looks like a frisbee. Assuming you remember that a *dax* is something you can play with, you could infer that it is probably the frisbee-like object. However, for the second time already, there is also a flower co-occurring with the noun *dax*. Indeed, you cannot take for granted that any of the objects in either of the two situations is the correct referent for *dax*. The girl in the first situation could have well just accidentally looked at the flowers and the man in the shop might have also looked for daxes that he could not find. That means that you are left with a conflict with regard to the meaning of *dax*: You may decide that it means (some kind of) flower, you may decide that it refers to (some kind of) frisbee, or you may also make no decision at all at this point because you are too uncertain.

There is evidence for the hypothesis that pieces of information which can only jointly result in word learning can still be helpful when presented separately: Arunachalam & Waxman (2010a) and Yuan & Fisher (2009) presented data suggesting that infant learners can still benefit from syntactic bootstrapping when the two co-cues linguistic information and visual information are not co-present: Infants first listened to small dialogues containing sentences with novel verbs within causative sentence frames (e.g., *The lady mooped my brother.*) without seeing a scene. Later in the experiment, they were presented two visually depicted situations (e.g., someone pushing someone else and two people waving) and were asked to identify the action matching the novel verb (*Where is mooping?*). Infants successfully used the information from sentence frames (i.e., that the verb is causative) to make up their mind. We will investigate whether the same is true for adults learners and for noun learning based on verbal constraints.

Some research has been conducted into the interplay of conflicting word learning cues (Moore et al., 1999; Houston-Price et al., 2006; Brandone et al., 2007). Houston-Price et al. (2006) manipulated two cues, gaze (someone looking at a potential referent) and salience (movement or activation of a potential referent) and studied about-2-year-old's noun-learning behavior. They found that infants neither used gaze information nor salience cues when both conflicted. Brandone et al. (2007) investigated the effect of salience (whether an action had an interesting result or not) and speaker information (i.e., social information/direct enacting and linguistic information/verb frame) on verb learning. They found that while 22-month-old infants did not learn verbs when speaker information was in conflict with the perceptual cue or when the perceptual cue was not informative, 34-month old children overrode perceptual information. While these results are interesting, they cannot straightforwardly be transferred to the behavior of adult learners: The way cues influence learners seems to underlie developmental changes and it is very likely that adult learners' mental capacities enable them to combine and weight different cues in another way than infants could do (see Gillette et al., 1999 versus Lidz et al., 2010). Moreover, while there are few studies (expect Experiments 1 to 3) which investigate the interplay of CSWL with other mechanisms (see Section 2.2.6), none addresses conflicting situations.

4.1. Experiment 4

4.1.1. Motivation

There were two major motivations for Experiment 4. The first was to investigate whether CSWL and SLCL are still helpful learning mechanisms in a setting which was further naturalized: With co-cues which were not co-present and with a more complex degree of referential uncertainty. The second motivation was to shed light on the role and nature of CSWL and SLCL by examining them when they support conflicting noun meanings.

We designed a learning setting in which each noun had two potential meanings (i.e., two potential referents, e.g., pizza and hat). One referent co-occurred with the noun in 83% of the noun's usages (*83% object* or *high-frequency object*, e.g., hat), the other referent co-occurred with the noun in only 50% of the cases (*50% object* or *low-frequency object*, e.g., pizza). The choice of exactly these two co-occurrence frequencies was determined by design characteristics described in the materials section. Additionally, each noun belonged to one of two conditions: In Condition ***N****(on restrictive)*, a noun was always preceded by a non-restrictive verb such as *take*. In Condition ***R****(estrictive)*, a noun was preceded by a restrictive verb such as *eat* half of the time. That means that in Condition N, learners' choice between the two potential referents was entirely based on co-occurrence frequencies (CSWL). We therefore hypothesized that learners would prefer to understand the 83% referent (hat) as the nouns meaning rather than the 50% referent (pizza). In Condition R, on the contrary, learners' choice about a noun's meaning was influenced by both co-occurrence frequencies (CSWL) and verbal restrictions (SLCL). Crucially, both CSWL and SLCL supported conflicting referents: CSWL supported the 83% referent (hat) whereas SLCL supported the 50% referent (pizza). Assuming that adults are capable of considering different cues in parallel even if defining their weight is non-trivial (see Sections 2.2.6 and above), we hypothesized, firstly, that CSWL and SLCL would both be used by word learners. Secondly, we hypothesized that both learning mechanisms would reduce each other's influence, leading into a mixture of final decisions. Given the highly associative nature of CSWL, we hypothesized that it may still influence learners on an attentive (unconscious) level even if SLCL is given priority.

Our design was characterized by the constraint that restrictive-verb information was never co-present with the objects which fulfilled the verbal constraints: In Condition R, the restrictive verbs (e.g., eat) were used in exactly that half of the trials which did not include the 50% referent (pizza). The first reason for integrating this manipulation was to avoid that the conflict between SLCL and CSWL would be too strong and direct and therefore too confusing for learners. The second objective was to examine whether learners can use verb information across situations. Based on the findings by Arunachalam & Waxman

(2010a) and Yuan & Fisher (2009) (see above), we hypothesized that learners would be able to transfer verb information across situations to later combine it with visual information in order to identify referents.

4.1.2. Methods

4.1.2.1. Participants

28 native speakers of German (students from different disciplines) who had not participated in either of Experiments 1 to 3 took part in Experiment 4. Four of these students had to be excluded due to unsuccessful verb learning (explained in Paragraph 4.1.2.2). Of the 24 remaining learners, four were males and 20 females. These 24 participants were aged between 20 and 35 years (mean age 24). They received €6 for taking part in the experiment.

4.1.2.2. Materials and Procedure

As in the first three experiments, Experiment 4 sought to teach participants a miniature semi-natural language. To make the imperfectness of sentence-scene relations more plausible and to help participants in perceiving the experiment as a more realistic situation, we asked them to imagine that they are visiting a friend in Artonesia and that they try to understand her friends and family who are speaking the local language of Artonesia (Artonesian). The language was similar to the one in the other experiments. It comprised two restrictive verbs ('eat' and 'sew') and two non-restrictive verbs ('take' and 'point at'), twelve nouns ('man', 'woman', and ten object names), and, as before, the article *si* (see Sections 3.1.2.2, 3.2.2.2, and 3.3.2.2).

The experiment consisted of the following phases: Verb Training 1, Verb Training 2, and verb testing, eye-tracker preparation, and verb repetition (Phase 1), noun learning Block 2 (Phase 2), vocabulary test Block 1 (Phase 3) and verb repetition, noun learning Block 2 (Phase 4), Vocabulary Test 2 (Phase 5), noun learning Block 3 (Phase 6), Vocabulary Test Block 3 (Phase 7) and final verb test. The reason for splitting noun learning and noun testing into blocks (Block 1, Block 2, Block 3) was to to facilitate learning. The entire experimental procedure lasted approximately 40 minutes.

Verb Learning In Phase 1, participants were familiarized with the verbs. Verb Training 1 consisted of a subpart of the materials of Experiment 3 and the procedure was similar (see Paragraph *Verb Learning* in Section 3.3.2.2). Participants were presented spoken verbs and simple animations concurrently. Instructions explicitly told them that the verbs are the names of the depicted actions and that they are asked to memorize these mappings. Each verb was repeated nine times (with two different animations per verb). To further facilitate verb familiarization, there was an additional forced-choice training (Verb Training 2). Pictures of the four actions were visible at the same time (the last position of the animations), one spoken verb was presented, and participants were requested to click onto the action matching the verb (each verb was tested twice). The verb testing was identical to that of Experiments 1, 2, and 3 (see Paragraphs *Verb Learning* in Sections 3.1.2.2, 3.2.2.2, and 3.3.2.2). Only participants who chose the correct actions for all four verbs were in this test were included into analyses. The eye-tracker was adjusted (see Paragraph *Verb Learning* in Section 3.1.2.2) and the verbs were quickly repeated (participants were shown the depictions and were prompted to name each verb twice).

Noun Learning Before noun learning Block 1 (Phase 2), participants were familiarized with the pictures of the objects to make sure that they could recognize all of them. The nouns for 'man' (*laki*) and 'woman' (*gadis*) were explicitly introduced, together with the depictions used in the experiment. We informed them that while scenes are often helpful for understanding the sentence, they do not necessarily fully correspond to the sentences. We considered this level of explicitness necessary to avoid confusion given the high referential uncertainty.

Materials for the noun-learning phases (Phases 2, 4, and 6) were basically as in the other experiments (see Paragraphs *Noun Learning* in Sections 3.1.2.2, 3.2.2.2, and 3.3.2.2; see Table 4.1): Participants were exposed to pairs of static scenes and spoken SVO-sentences. These sentences consisted of the verbs that had been learned before and novel nouns as subjects and objects. Scenes depicted semi-natural indoor-scenes with characters and objects. Participants' task was to learn the noun meanings.

In contrast to Experiments 1 to 3, however, each of the ten object nouns to be learned had two potential meanings, which means that there were two referents per noun, one referring to a food item and one referring to a clothing item (e.g., socks and corn for the noun *daram* and dress and apple for the noun *firel*, see Table 4.1). Therefore, 20 different objects were depicted on all trial scenes (ten food objects and ten clothing objects). One of the two meanings for each noun was *CSWL-supported*: The co-occurrence of the noun with that object was 83% (e.g., socks for *daram*). The other meaning was less supported by CSWL (co-occurrence only 50%, e.g., corn for *daram*) but still more supported than the other distractors which co-occurred only 17% with a noun (e.g., jacket for *daram*). Given that each noun was presented six times, the 83% object (e.g., socks) co-occurred in five of these six trials, the 50% object (e.g., corn) co-occurred three times and each 17% distractor (e.g., jumper) co-occurred only one time. We chose exactly these three co-occurrence frequencies (17%, 50%, 83%) because they had to be sufficiently different from each other and the high-frequency object was supposed to not co-occur 100%.

We additionally introduced the within-participant two-level factor Verb Type: In Con-

Table 4.1.: Example trials Exp. 4

Verb Type	Cat.	Type	Sentence	Scene
Non-restrictive	1	1	*Si laki gumbumema si daram* 'The man will take the DARAM.'	socks (83%), corn (50%), dress (17%)
		2	*Si gadis gumbumema si daram* 'The woman will take the DARAM.'	woman, socks (83%), corn (50%), top (17%)
		3	*Si laki gumbumema si daram* 'The man will take the DARAM.'	socks (83%), corn (50%), pizza (17%)
	2	4	*Si gadis tambamema si daram* 'The woman will take the DARAM.'	woman, socks (83%), jacket (17%), jumper (17%)
		5	*Si gadis tambamema si daram* 'The woman will take the DARAM.'	woman, socks (83%), skirt 17%)
	3	6	*Si laki tambamema si daram* 'The man will take the DARAM.'	man
Restrictive	1	1	*Si laki gumbumema si firel* 'The man will take the FIREL.'	dress (83%), apple (50%), socks (17%)
		2	*Si gadis gumbumema si firel* 'The woman will take the FIREL.'	woman, dress (83%), apple (50%), top (17%)
		3	*Si laki gumbumema si firel* 'The man will take the FIREL.'	dress (83%), apple (50%), pizza (17%)
	2	4	*Si gadis bermamema si firel* 'The woman will eat the FIREL.'	woman, dress (83%), jacket (17%), jumper (17%)
		5	*Si gadis bermamema si firel* 'The woman will eat the FIREL.'	woman, dress (83%), skirt (17%)
	3	6	*Si laki bermamema si firel* 'The man will eat the FIREL.'	man

dition ***N****(on-restrictive)*, nouns always followed non-restrictive verbs (4.1). That means that while the 83% object was CSWL-supported, the 50% object was not supported by any other cue. In Condition ***R****(estrictive)*, however, nouns occurred with a restrictive verb in 50% of the trials (see Table 4.1, Trials 4-6). Crucially, these restrictive verbs supported the 50% referents which means that the 50% referents were *SLCL-supported*. The hypothesis that apple is the meaning of the noun *firel*, for instance, was SLCL-supported because *firel* was sometimes preceded by the verb bermamema ('to eat') and an apple is edible. That means that, in Condition R, one meaning was CSWL-supported (dress) whereas the other meaning was SLCL-supported (apple).

Each object was assigned to one noun only, that means that no object was a low-frequency candidate for one noun and a high-frequency candidate for another noun. This control was important to guarantee that mutual exclusivity could not help learners to decide between one of the two meanings.

Given that nouns' referents were sometimes not depicted, referential uncertainty was made more complex for the learner, in order to make the learning setting more natural: Not everything in the scene was mentioned and not everything mentioned was in the scene. Additionally crucial was that restrictive verbs were used only in that half of the trials in which the scene *did not* include the SLCL-supported referent. That means that restrictive verbs and referents matching the verb's semantic category were never co-present and verb information had to be memorized across trials. We used this manipulation to avoid that the conflict between SLCL and CSWL was too direct and therefore confusing for learners. Moreover, we aimed at investigating whether learners are able to use verb information across trials (as in Arunachalam & Waxman, 2010a and Yuan &Fisher, 2009).

Trials (i.e., scene-sentence-pairs) belonged to three different *trial categories* which varied in scenes and scene-noun relations. This was necessary to ensure that the high-frequency referent and the low-frequency referent had different co-occurrence frequencies (83% and 50%) and that restrictive verbs and matching objects did not co-occur. Trials of Category 1 (three of six trials) were characterized, firstly, by a scene that contained the low-frequency referent and the high-frequency referent and, secondly, by the fact that the verb was always non-restrictive in both conditions. Trials of Category 2 (two of six trials) contained scenes depicting the high-frequency referent but not the low-frequency referent and sentences with restrictive verbs in Condition R. Trials of Category 3 (one of six trials) were characterized by verbs which were restrictive in Condition R and by scenes with an agent depicted only. These trials were integrated to implicitly remind participants that the relation between sentences and scenes does not always have to be perfect in that not all referents are depicted.

There were further within-category differences for trials in Trial Categories 1 and 2 (which I refer to as trial *types*). These differences were made in order to increase the variability and therefore naturalness of scene-sentence relations (and to repeat and distribute all objects equally over the course of the entire experiment). The difference between Trials of Type 1, 2, and 3 (all Trial Category 1) was the distractor type and the character depiction: The scene in trials of Type 1 depicted the high-frequency object, the low-frequency object, and a distractor of the high-frequency category. Scenes in trials of Type 2 additionally depicted a character. Finally, Trial-Type 3 scenes embedded the high-frequency object and the low-frequency object and a distractor from the low-frequency category. The difference between trials of Type 4 and 5 (both Trial Category 2) was the number of distractors: While both scenes contained a character, the high-frequency object, and a distractor from the high-frequency category, in Type 5 a second distractor from the high-frequency category was additionally depicted.

Participants were presented trials according to one of four lists: There were two world-word-mappings and two assignments of items to conditions. By assigning the items to both conditions across lists, the assignment to object type (low-frequency or high-frequency) was manipulated at the same time. The noun *firel*, for instance, was in Condition R and had a food object as low-frequency referent (and a clothing item as high-frequency object) in two of the lists. In the two remaining lists, it belonged to Condition N and had a clothing item as low-frequency referent (and a food item as high-frequency candidate).

In noun-learning Block 1 (Phase 2), two nouns were introduced, one food item and one clothing item (12 trials). The condition of the two nouns depended on the participant's list but it was always the case that one noun was in Condition N, the other one in Condition R. Block 2 (Phase 4) and Block 3 (Phase 6) each contained four novel nouns (24 trials per block), one clothing item of Condition N, one food item of Condition N, one clothing item of Condition R, and one food item of Condition R.

The presentation of the 60 trials was pseudo-randomized within blocks with the following constraints: Each trial of Trial Categories 2 and 3 was directly followed by a trial of Trial Category 1. That means that each trial containing a restrictive verb (e.g., eat) was directly followed by a trial containing a referent that matched this verb's constraints (e.g., corn). The reason for this control was to reduce the difficulty of using verb information across trials. The resulting trial pairs were then randomized in order.

Vocabulary Tests As in the other experiments, participants' task in the vocabulary tests (Phases 3, 5, and 7) was to make a forced choice about the meanings of nouns: They were asked to click onto the object depicted on the scene which matched the spoken noun. Opposed to Experiments 1 to 3, all 20 objects were shown for each test trial. This modification addresses recent criticism of the standard CSWL tests: Smith et al. (2009) argue that a constrained set of objects in the test simplifies the task and cannot be used to evaluate pure CSWL (see Section 2.2.1). Confidence ratings were recorded, as in Experiment

3.

In vocabulary test Block 1 (Phase 3), we tested knowledge about the two nouns which participants had been familiarized with in noun learning Block 1 (Phase 2); in the vocabulary test Block 2 (Phase 5), the four nouns presented in noun-learning Block 2 (Phase 4) were assessed; in the vocabulary test Block 3 (Phase 7), participants were tested on their knowledge about the four nouns presented in noun-learning Block 3 (Phase 6), accordingly.

The vocabulary test in Block 3 (Phase 7) was followed by another verb test: Participants were asked to forced-choice decide which verb matched which action depiction. This was done in order to assess whether participants were entirely familiar with all four verbs at the end of the experiment, which was required to be included into analyses.

4.1.3. Predictions

Based on the evidence by Arunachalam & Waxman (2010a) that adults are able to use co-cues across trials (see beginning of Chapter 4), we hypothesized that to learn nouns in Condition R, word learners would use SLCL, which supported the 50% target. Additionally we hypothesized that noun learning in both conditions would be influenced by CSWL, which supported the 83% target (see Section 4.1.1). Both hypotheses together predicted, firstly, a clear preference in Condition N to select the 83% target in the vocabulary tests (Phases 3, 5, and 7) and, secondly, a conflict between choosing the SLCL-supported 50% target and the CSWL-supported 83% target in Condition R. We expected that this conflict would result in a mixture of choices in vocabulary test trials of Condition R with about as many 83%-choices as 50%-choices.

Our hypotheses moreover motivate some predictions regarding eye-movements during noun learning (Phases 2, 4, and 6) and noun testing (Phases 3, 5, and 7). Specifically, hypothesizing that learners in Condition N are only influenced by CSWL, which supported the 83% object, we expected a clear preference for inspecting this object during learning and testing. On the contrary, for Condition R, we predicted participants to look more often at both the 83% object and the 50% object than at distractor objects, hypothesizing learners to use CSWL (supporting the 83% object) and SLCL (supporting the 50% object). More specifically, we expected that participants would either inspect both objects equally often or inspect the 83% object more than the 50% object. The first possibility would be supported by the hypothesis that CSWL and SLCL are equally effective not only off-line but also on-line. The second possibility would be supported by the hypothesis that CSWL is more influential than SLCL on an associative and unconscious level which may be reflected by eye-movements.

4.1.4. Data Analysis and Results

4.1.4.1. Off-line Results

Verb learning as assessed by the verb test in Phase 1 and the final verb test was perfect in all 24 participants whose results are reported below.

Performance in noun learning (i.e., learning either the low-frequency or the high-frequency meaning) was clearly better than chance (10%): 84.2% for both conditions taken together ($t(23) = 26.319, p < .001$), 87.5% for Condition N ($t(23) = 24.665, p < .001$), and 80.8% in Condition R ($t(23) = 20.206, p < .001$). To analyze whether factor Verb Type has an effect on the chosen meaning in the vocabulary test, we used linear mixed-effects models, using logistic regression for the categorical learning rates (logit link function), with participant and item as random factors. The two models including and excluding factor Verb Type were moreover compared with a Chi-Square Test (see Section 3.1.4.2 for further explanation about this way of analyzing). In particular, we only analyzed the subset of choices which were correct (either the high-frequency object or the low-frequency object). Importantly, we found a main effect of Verb Type for the chosen meaning ($\chi(1) = 59.300, p < .001$, Table 4.2): In N, learners chose the high-frequency meaning 97% of the times and the low frequency meaning only 3%. in Condition R, however, high- and low-frequency meanings were chosen about equally often (high: 48%, low: 52%) (see Figure 4.1). The effect of Verb Type on chosen meaning also confirms that learners were able to use verb information across trials. The average confidence rating was 6.9 and there was no difference between conditions (6.8 in Condition N and 7.0 in Condition R). There was a positive correlation between confidence ratings and learning rates ($r_s = .342, p < .001$).

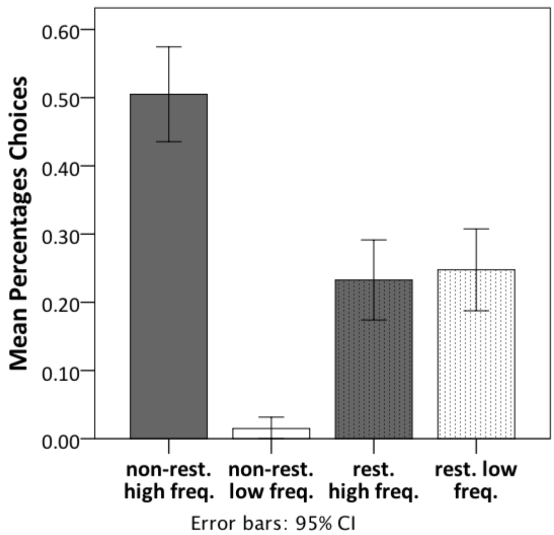


Figure 4.1.: Percentages of vocabulary-test choices (high-frequency object / low-frequency object) in verb types, Exp. 4

Table 4.2.: Lmer models for chosen meaning (low-frequency/high-frequency object), Exp. 4
$chosenmeaning \sim VerbType + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
1	(Int) (Non-restrictive)	−3.5781	0.601	−5.950	< .001
2	Restrictive	3.425	0.532	6.442	< .001

4.1.4.2. On-line Results

Eye-movements during Verb and NP2 in Trials of Trial Category 1 Eye-movement data from noun-learning trials was analyzed similarly as in Experiments 1 to 3 (see Section 3.1.4.2): We conducted logistic regression analyses by entering the binomial data (inspection or no inspection at certain time to a specific ROI) into linear mixed effects (*lmer*) models with a logit link function (from the lme4 package in R, Bates, 2005). However, in Experiment 4, Block was added as a random factor (in addition to Participants and Items): While we were not interested in block differences, we expected that eye-movements may change over the course of the experiment (for instance due to learning of noun meaning). For continuous eye-movement data (frequency of inspections to either ROI and inspections during vocabulary-test trials), we used linear regression analyses and calculated Monte Carlo Markov Chain values (MCMCs) (Baayen et al., 2008).

We firstly conducted analyses of gaze to the two ROIs low-frequency object and high-frequency object during verb and NP2 for trials of Category 1 (recall, verbs were non-restrictive here in both conditions but the preceding trial contained a restrictive verb for R-trials; see Table 4.4, Rows $1-8$). Looks to the character and to distractors could not be included because none of them was depicted in all three trial types included in Trial Category 1 (see Table 4.1). See Figure 4.2 for timegraphs.

During verb region, the high-frequency object was inspected more often than the low-frequency object (main effect ROI both conditions: $\chi(1) = 10.29, p < .01$). This effect was significant in Condition N ($\chi(1) = 11.700, p < .001$; see Table 4.4, Row 2) but not in Condition R ($\chi(1) = 1.358, p = .244$; see Table 4.4, Row 4). See barchart in Figure 4.3. We neither found an effect of Verb Type ($\chi(1) = 0.300, p = .584$) nor an interaction of ROI and Verb Type ($\chi(2) = 3.217, p = .200$).

When NP2 was spoken (from onset of the noun on), learners looked significantly more often at the high-frequency object than the low-frequency object in both conditions (main effect ROI both conditions: $\chi(1) = 38.549, p < .001$). While this effect was clear in Condition N ($\chi(1) = 52.104, p < .001$; see Table 4.4, Row 6), it was only marginal in

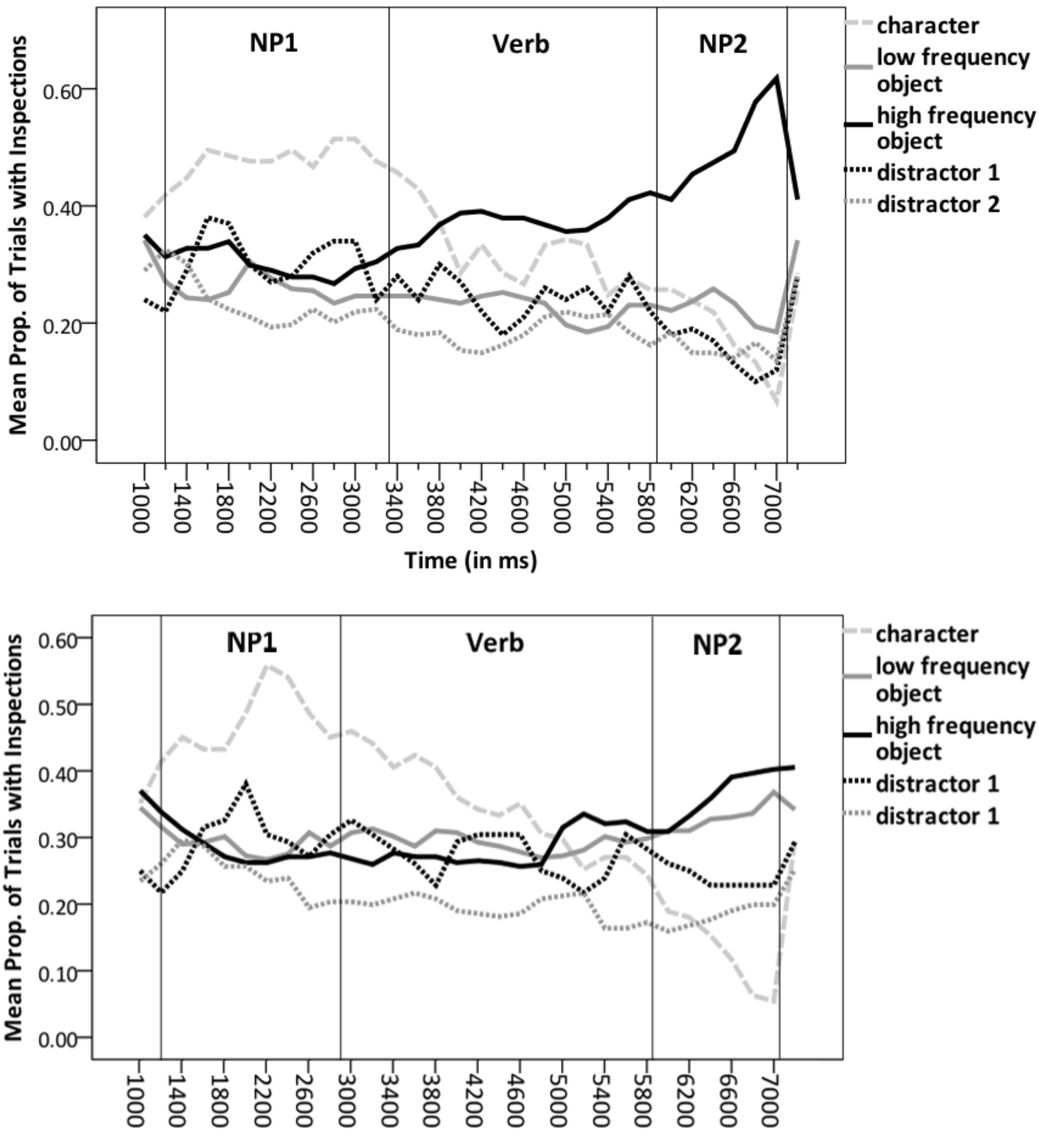


Figure 4.2.: Timegraph Experiment 4, Trial Category 1, Condition N (top) and Condition R (bottom)

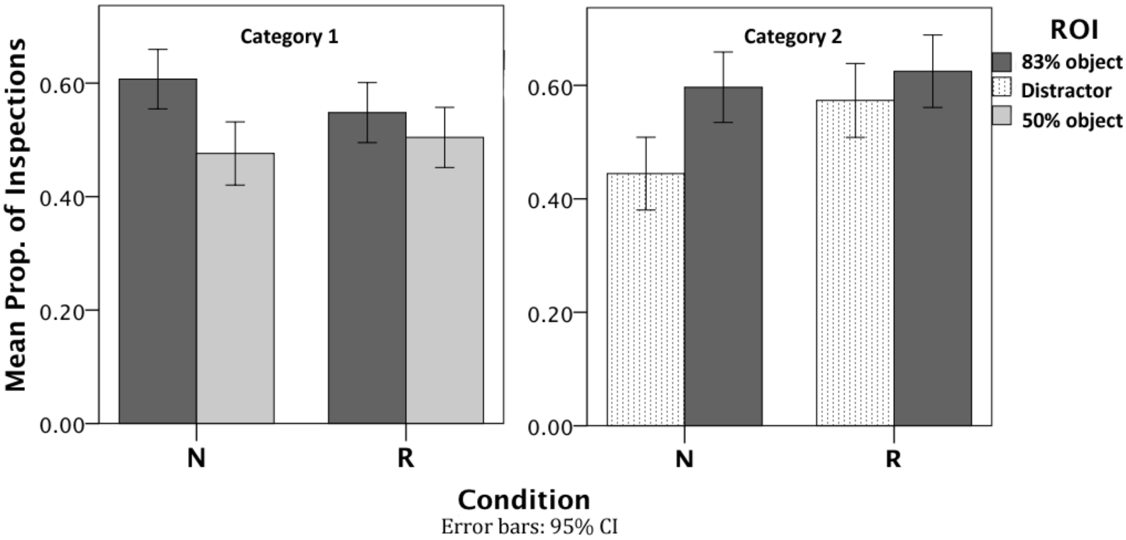


Figure 4.3.: Proportions of trial with at least one inspection to ROIs during the verb, Trial Categories 1 (left chart) and 2 (right chart), Exp. 4

Table 4.3.: Lmer models and p-values from MCMC sampling for number of inspections to any ROI, depending on meaning chosen, during learning trials of Category 1, Condition R, Exp. 4
$InspectionsAnyROI \sim meaningchosen + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
1	(Int.) (low-freq.)	10.867	0.678	16.018	10.871	$< .001$	< 001
2	high-freq.	-4.433	0.960	-4.621	-4.437	$< .001$	$< .001$

Condition R ($\chi(1) = 3.015, p = .083$, see Table 4.4, Row 8). We found no effect of condition ($\chi(1) = 0.905, p = .342$). However, there was an interaction between factors ROI and Condition ($\chi(2) = 16.773, p < .001$): The difference between looks to both candidates was larger in Condition N than in Condition R (see Figure 4.4).

We additionally analyzed whether the amount of inspections to any ROI during R-trials of Trial Category 1 (averaged across participants) was different depending on the meaning that was chosen in the vocabulary test. Given that the dependent variable was continuous, linear regression analyses rather than logistic regression analyses were conducted. We found that more looks are made to any region when the high-frequency object is chosen then when the low-frequency object is chosen ($\chi(1) = 18.807, p < .001$; Table 4.3). That means that participants inspected scenes during learning more at length when the object that they chose in the final test was the high-frequency target than when it was the low-frequency target.

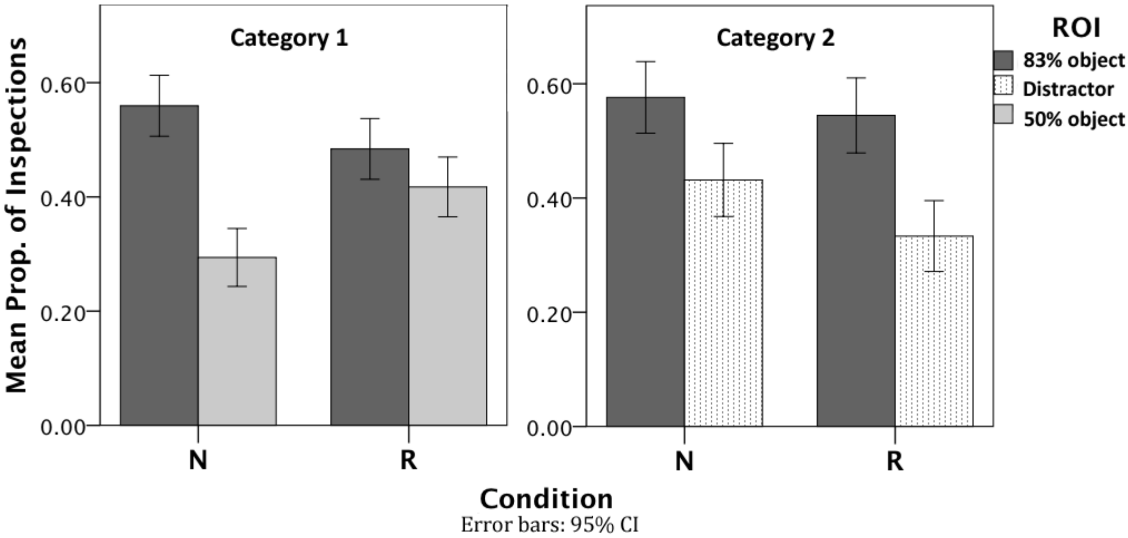


Figure 4.4.: Proportions of trial with at least one inspection to ROIs during NP2, Trial Categories 1 (left chart) and 2 (right chart), Exp. 4

Table 4.4.: Lmer models for inspections on low-frequency object vs. high-frequency object during verb and NP2, Trial Categories 1 and 2, Conditions N and R, Exp. 4
$Inspections \sim ROI + (1|sub) + (1|item) + (1|blocks), family = binomial(link = "logit")$

		Predictor	Coef.	SE	Wald z	p
Trial Category 1						
1	*N,V*	(Int) (high-f.)	−0.106	0.130	−0.816	$= .041$
2		low-f.	0.549	0.160	3.431	$< .001$
3	*R,V*	(Int) (high-f.)	0.011	0.138	0.077	$= .939$
4		low-f.	0.182	0.154	1.180	$= .238$
5	*N,NP2*	(Int) (high-f.)	−0.987	0.180	−5.492	$< .001$
6		low-f.	1.222	0.171	7.136	$< .001$
7	*R,NP2*	(Int) (high-f.)	−0.351	0.145	−2.422	$< .050$
8		low-f.	0.273	0.156	1.754	$= .079$
Trial Category 2						
9	*N,V*	(Int) (high-f.)	0.444	0.193	2.302	$< .050$
10		low-f.	−0.658	0.188	−3.511	$< .001$
11	*R,V*	(Int) (high-f.)	0.623	0.301	2.070	$< .050$
12		low-f.	−2.235	0.198	−1.189	$= .235$
13	*N,NP2*	(Int) (high-f.)	0.334	0.210	1.587	$= .112$
14		low-f.	−0.672	0.193	−3.481	$< .001$
15	*R,NP2*	(Int) (high-f.)	0.250	0.286	0.874	$= .382$
16		low-f.	−0.983	0.203	−4.833	$< .001$

Eye-movements during Verb and NP2 in Trials of Trial Category 2 Looks during trials of Category 2 were also analyzed (high-frequency but not low-frequency object included, restrictive verb in Condition R, non-restrictive verb in Condition N; see Table 4.4, lines $9-16$). Looks to one distractor (present in both trial types) was included as baseline. See timegraphs in Figure 4.5.

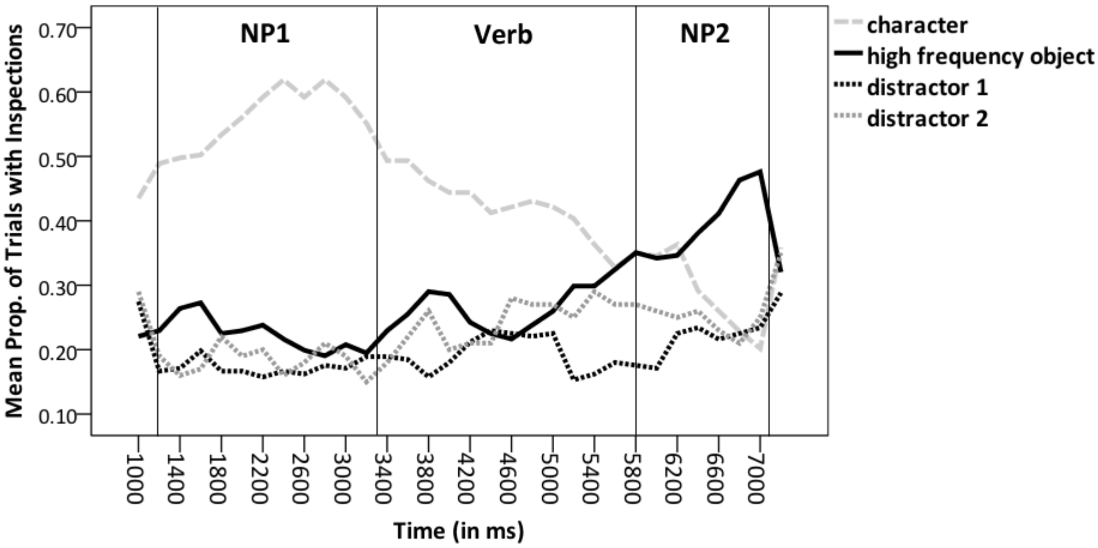
For verb region, we found an effect of ROI (more looks to the high-frequency object: $\chi(1) = 1.369, p = .242$), an effect of Verb Type (more looks in Condition R than N: $\chi(1) = 6.067, p < .05$), but no interaction ($\chi(2) = 2.619, p = .106$; see Figure 4.3). For NP2, we found an effect of ROI (more looks to the high-frequency object: $\chi(1) = 33.001, p < .001$), an effect of Verb Type (more looks in Condition N than R: $\chi(1) = 4.077, p < .05$), and no interaction ($\chi(1) = 1.009, p = .315$; see Figure 4.4. In Condition N, the high-frequency object was inspected significantly more often than the distractor in both verb region ($\chi(1) = 12.082, p < .001$; Table 4.4, Row 10) and NP2 ($\chi(1) = 11.626, p < .001$; Table 4.4, Row 14). In Condition R, there was no effect in verb region ($\chi(1) = 1.358, p = .244$; Table 4.4, Row 12) but learners looked significantly more often to the high-frequency object than to the distractor object during NP2 ($\chi(1) = 23.046, p < .001$; Table 4.4, Row 16).

Eye-movements during Vocabulary Testing We further analyzed the number of inspections to the high-frequency candidate and the low-frequency candidate during vocabulary testing, averaged across participants (inspections to ROIs from speech onset, 1000ms after trials start, data averaged across trials, see Figure 4.6). The reason for using this dependent variable was that analyzing whether ROIs were inspected or not in a trial was not informative due to the large time region. Given that it was continuous, linear regression analyses rather than logistic regression analyses were conducted. While we found that the high-frequency object was inspected significantly more often than the low-frequency object in both conditions (both: $\chi(1) = 54.778, p < .001$; N: $\chi(1) = 41.657, p < .001$; R: $\chi(1) = 23.690, p < .001$, see Table 4.5), there was also a significant interaction between factors ROI and Condition ($\chi(2) = 20.781, p < .001$): The difference between looks to both objects was much greater in Condition N than Condition R. There was no effect of Verb Type (not more looks in one or the other: ($\chi(1) = 1.256, p = .262$).

In Condition R, there were more looks (to any ROI) when the high-frequency meaning was chosen than when the low-frequency object was chosen. This difference did not reach statistical significance ($\chi(1) = 0.916, p = .339$).

4.1.5. Summary and Discussion

To summarize the results of Experiment 4, learning rates for both conditions were clearly above chance. While learners almost always selected the high-freqeuncy target in vocabulary-



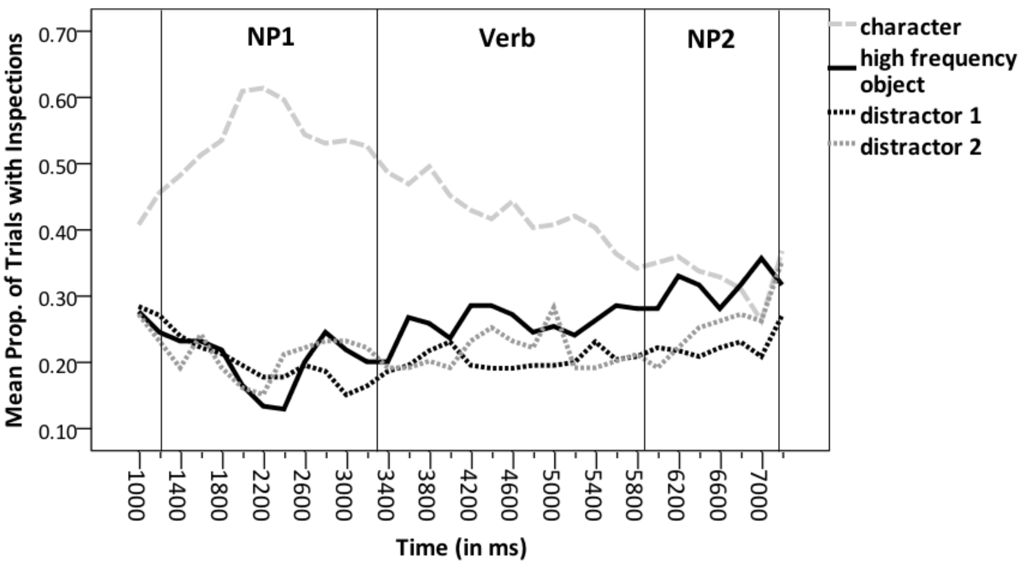


Figure 4.5.: Timegraph Experiment 4, Trial Category 2, Condition N (top) and Condition R (bottom)

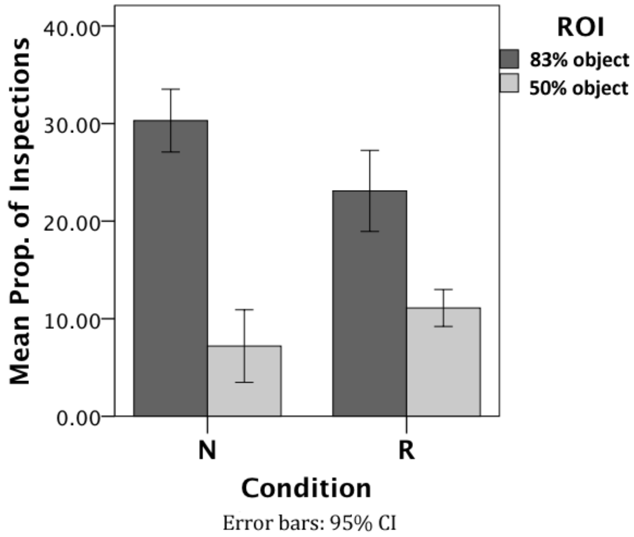


Figure 4.6.: Proportions of trials with at least one inspection to ROIs during vocabulary test (from speech onset, 1000ms after trials start), Exp. 4

Table 4.5.: Lmer models & p-values from MCMC sampling for inspections to low-frequency object vs. high-frequency object during vocabulary test, Exp. 4
$inspections \sim ROI + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
Condition N							
1	(Int.) (low-freq.)	7.200	1.538	4.682	7.210	$< .001$	$< .001$
2	high-freq.	23.100	1.650	14.004	23.07	$< .001$	$< .001$
Condition R							
3	(Int.) (low-freq.)	11.100	1.426	7.787	11.110	$< .001$	$< .001$
4	high-freq.	12.000	1.549	7.746	11.990	$< .001$	$< .001$

test trials of Condition N, they decided to choose the low-frequency target and the high-frequency target equally often in Condition R. Eye-movements during both learning and testing reveal a bias to inspect the high-frequency target in both conditions: In trials of Category 1, it was inspected more often than the low-frequency target and in Category 2 it was looked at more frequently than the distractor. This means that even when the low-frequency candidate was chosen in vocabulary test trials of Condition R, there was a tendency for looks to the high-frequency candidate during learning and testing. Moreover, in R-trials during learning and testing, participants looked around from region to region more frequently when the high-frequency object was chosen than when the low-frequency object was chosen. This suggests that conducting CSWL required considering all objects in parallel to find the highest co-occurrences whereas this was less necessary for learning via SLCL. Eye-movements during the verb region in learning trials of both categories reveal differences in conditions: Whereas no effect was found in R-trials, in Condition N, unexpectedly, the high-frequency object was looked at significantly more often than both the low-frequency referent in trials of Category 1 and the distractor in trials of Category 2.

Learners' decisions in the vocabulary test revealed a clear difference of condition: While the high-frequency object was unambiguously favored in Condition N, the high-frequency object (not supported by the verb) and the low-frequency object (supported by the verb) were chosen equally often in Condition R. This shows that verbs (SLCL) and statistics (CSWL) had a similar impact on vocabulary decision, with verb information overriding cross-situational statistical information half of the time. Eye movements during training and testing confirm that the influence of statistics on learner's on-line behavior in Condition R was modulated by the verb (which was presented in the trials directly preceding): The differences between the number of inspections on high-frequency objects and low-frequency objects were reliably larger in Condition N than Condition R. Interestingly, the fact that the high-frequency target was still looked at reliably more often than the low-frequency object in both conditions reveals that statistics were not completely ignored neither during learning nor during testing even when SLCL was given priority in the final decision. Furthermore, our findings demonstrate that learners made use of verbal restrictions across trials (in accordance with the results of Arunachalam & Waxman, 2010a).

We speculate that the impact of verbs may be even stronger in natural settings in general and when verbs are more deeply integrated into learners' mental lexicon. It is likely the case, firstly, that people are more tolerant to implausibility during an experimental situation and, secondly, that they do not perfectly rely on their knowledge about something that they have just learned.

One could argue that it is surprising that we found significantly more looks to the high-frequency object than to the low-frequency object in the verb region of N-trials of

Trial Category 1. However, learners may have learned over the course of the experiment that there were trial pairs (i.e., that each trial of Categories 2 and 3 was preceded by a trial of Category 1). Specifically, it is possible that participants learned that each noun was repeated in the following trial and that scenes of both of the two trials contained the high-frequency referent. They may have then used this information to identify the high-frequency referent in the second of the two trials even before it was named. Although the same prediction effect could have emerged for Condition R, it is not the case that the high-frequency referent was inspected significantly more often than the low frequency object during the verb region of R-trials of Trial Category 1. Rather, participants looked at the the 50% object, which matched the verbal constraints of the trial before, about equally often as at the 83% object. This indicates that learners in Condition R were influenced by both SLCL, supporting the 83% referent, and CSWL, supporting the 50% referent.

Gaze results during trials of Category 2 are broadly as expected but less crucial. During NP2 participants tended to look at the high-frequency object in both conditions, which is plausible due to learning over time. There were reliably more inspections to any ROI in Condition N than Condition R: In Condition R, learners often did not inspect any regions on the scene probably because they were aware of the fact that no object matched the verbal restriction. The fact that the high-frequency object was inspected more often than the distractor object during the verb in Condition N, however, is surprising because it was neither the case that the verb provided any useful information nor was there a learnable cue in trial order. In Condition R, there was no effect of ROI in verb region which is plausible because there was no object on the scene which matched the verbal restrictions. There were more looks to any ROI in verb region for Condition R than Condition N which might be attributed to learners searching for a semantic match to the verb (which was not present).

Another finding comes from the relation between chosen meanings in the vocabulary test and eye movements during noun learning trials of Category 1, in Condition R: Participants made more inspections (to any ROI) during NP2 when they chose the 83% meaning in the test later then when they chose the 50% meaning. That is, learners who chose the target based on CSWL made more inspections on all objects in the scene than learners who were guided by SLCL towards the low-frequency target. This may reflect that identifying referents based on SLCL is more deterministic than identifying referents based on CSWL: While CSWL learning involves considering and memorizing all objects in order to decide which one co-occurs most often with a noun, SLCL learning does not. We found the same tendency for the relation between chosen meanings and gaze during vocabulary-test trials in Condition R, indicating that learners may have also made their final choice on a more deterministic level for verb-supported meanings (the 50% candidate) than CSWL-supported meanings (the 83% candidate).

It has to be taken into account for the interpretation of eye movements that the number of trials that could be analyzed for different Trial-Type Categories and different Verb Types was not very high (Trial Category 1: 15 per condition; Trial Category 2: 10 per condition). This might have caused missing significances on the one hand and artifacts on the other hand. Decisions that learners made in the vocabulary test, however, offer a solid basis for interpretation.

Taken together, the most important findings coming from Experiment 4 are the following: When SLCL and CSWL are in conflict, both have an equally strong effect on noun learning (decisions in the vocabulary test). However, CSWL was not completely ignored even when SLCL was given priority (eye-movements during learning trials of Category 1). Moreover, learning based on SLCL may be more deterministic than learning based on CSWL.

4.2. Interim Summary

Findings from Experiment 4 clearly reveal that, when CSWL and SLCL are in conflict, both are about equally influential with regard to final decisions in a forced-choice vocabulary test. However, as eye-movements suggest, even when SLCL is given priority, CSWL is still not completely ignored neither during learning nor during testing. While these findings, together with those from Experiments 1 to 3, shed light on the interplay of SLCL and CSWL, final conclusions can not yet been drawn. In particular, given that SLCL (as defined in our experiments) helps learners to identify referents in a clearly more direct way than CSWL, it is somewhat surprising that we have not yet found evidence that SLCL can take priority over CSWL. One potential situation which could support this kind of interplay is when cues are not conflicting but just redundantly co-present, that is, when CSWL and SLCL are independently applicable.

Eye-movement data from Experiment 4 suggests that CSWL is a less deterministic way of learning than SLCL and that it is based on relatively low-level associations: During learning and testing, participants looked around from region to region more frequently when they chose the CSWL-supported high-frequency object in the test than when they chose the SLCL-supported low-frequency object, indicating that they generally considered everything when conducting CSWL. Given that this measurement of the 'overall number of inspections to any region of interest' is rather unconventional, our interpretation is only preliminary and needs to be verified.

Finally, given that the vocabulary test in Experiment 4, as the tests in Experiments 1 to 3, takes place immediately after training, another gap needs to be closed to finish this investigation: It needs to be examined whether CSWL and SLCL also lead into deeper learning than can be assessed with a forced-choice vocabulary test (and sentence judgement test) after 30 minutes of exposure to the novel language.

5. The Interaction of CSWL and SLCL

Experiments 1 to 4 provide important evidence regarding the interplay of CSWL and SLCL (i.e., learning based on verbal constraints): Firstly, CSWL and SLCL are powerful word-learning mechanisms for adults, even in naturalized experimental settings; secondly, SLCL enables the rapid on-line prediction and identification of referents; thirdly, both mechanisms operate jointly or complementarily to support learning of word meanings; fourthly, both mechanisms are approximately equally influential when in conflict, with a slight dominance of CSWL (in our setting). It remains unclear, however, how precisely both mechanisms work and why they interact the way they do. Furthermore, so far we have only tested very short-term learning in that the assessment took part directly after training. It is an important question whether participants' command of word meanings lasts longer than the experimental session and whether it underlies consolidation changes (Lindsay & Gaskell, 2010).

In the final study (Experiment 5), we therefore studied the underlying nature of CSWL and SLCL mechanisms, on one hand, and vocabulary retention one day after training, on other hand. In particular, the potential parallelism and probabilism of CSWL was examined by testing learners' sensibility to different co-occurrence frequencies. Moreover, we designed a learning setting in which CSWL and SLCL are independently applicable (i.e., redundantly available) in order to investigate whether the hypothesized determinism of SLCL blocks the use of CSWL. To test learners' vocabulary retention, we repeated the vocabulary test about 24 hours after training. To more specifically motivate Experiment 5, we will now briefly review findings from both recent studies and our own experiments concerning the nature of CSWL and SLCL and the role of consolidation in word learning.

5.1. The Nature of CSWL and SLCL

CSWL is incremental in that learning takes place over the course of multiple trials, stepwise and over time (Yu & Smith, 2007; Blythe et al., 2010). This implies that there should be different stages of fragmentary knowledge before the whole meaning of a word is established. Yurovsky, Fricker, Yu & Smith (2010) documented that this is indeed the case and that partial knowledge of a world-word mapping not only facilitates learning this particular mapping but also boosts learning of other words: In their experiment, learners were first exposed to a block of CSWL-training and a forced-choice vocabulary test. Those world-word mappings which were not correctly identified in the test were then used again in a second block of CSWL-training, together with entirely unfamiliar world-word mappings. Data from the forced-choice vocabulary test after Block 2 revealed that learning rates for the mappings which had been already presented in Block 1 were significantly higher than learning rates for other words. Interestingly, learning rates for all other mappings in Block 2 were also increased (compared to learning rates of participants who had not taken part in Block 1).

CSWL is moreover a parallel way of learning because one word is temporarily associated with different referents at the same time. This means that more than one potential meaning (i.e., world-word mapping) is mentally stored. These different mappings are then weighted across trials. CSWL can also be considered probabilistic, in that each mapping has a probability to be the correct one, or in other words, each potential referent linked to the word has a certain probability to be the correct referent. That means that the probability is determined by co-occurrence frequencies of world-word relations. Evidence for the last two claims comes from Vouloumanos (2008) who presented results revealing that CSWL-users are influenced by fine-grained differences in co-occurrence frequencies: In a forced-choice vocabulary test with two choices, participants were able to differentiate between co-occurrence frequencies of 0% and 10%, 0% and 20%, 0% and 60%, 0% and 80%, 0% and 100%, 10% and 20%, 10% and 60%, as well as 10% and 80%.

The findings of Yurovsky et al. and Vouloumanos are compelling in that they change the picture of word learning that has been dominant for some time: They suggest that word knowledge is not generally a binary state but can be an accumulation of incomplete hypotheses that are revised over time. This might be crucial for word learners since perfectly disambiguating situations are rare in realistic settings which makes it necessary to collect pieces of evidence incrementally. However, given that both studies use highly simplified learning settings, it is an important enterprise to further evaluate the generality of their findings.

Compared to CSWL, SLCL offers a more direct and deterministic learning cue because it often helps learners to identify referents immediately and unambiguously (see Altmann and Kamide, 1999). Eye-movement data from Experiments 1 to 4 support this hypothesis: When verb information was restrictive (Experiment 1, Condition SVO of Experiment 2, and Conditions NoRU and LowRU of Experiment 3), only objects fulfilling these constraints were considered. Determinism in SLCL is also reflected by the overall distribution of inspections in Condition R of Experiment 4: Learners looked around at other scene objects less often when they chose the verb-supported meaning in the vocabulary test than when they picked the CSWL-supported meaning. Verb-driven learning also turned out to be a very helpful cue in our experiments: On-line identification of referents which were selected by a verb was quick and successful (Experiment 1, Condition SVO of Experiment 2, and Conditions NoRU and LowRU of Experiment 3), as evidenced by eye-movements. Furthermore, learning rates and confidence ratings are higher for those nouns denoting these referents than learning rates for nouns learned via CSWL (Condition NoRU versus Condition HighRU of Experiment 3). This is probably due to, firstly, second-language learners' general experience in exploiting linguistic contexts to make inferences when they encounter knowledge gaps and, secondly, to people's rich knowledge about plausible states and relations in the world (such as 'only edible things are likely to be eaten').

There is another potentially interesting characteristic in the nature of SLCL. Given that the selection of potential referents in SLCL is based on category membership, it is likely that SLCL enhances learner's sensitivity for category associations. That means that SLCL-learned words are more likely to be associated with a category (e.g., food) additional to a particular meaning (e.g., banana) than CSWL-learned words.

5.2. Consolidation in Word Learning

Newly acquired knowledge and skills need to be consolidated - integrated into memory and generalized. Sleep has been shown to play an important role for this process in several cognitive areas (Wagner, Gais, Haider, Verleger & Born, 2004) and specifically for language skills (Fenn et al., 2003; Gómez, Bootzin & Nadel, 2006; Dumay & Gaskell, 2007). Recent studies reveal that sleep improves the generalization of phonological categories (Fenn, Nusbaum & Margoliash, 2003; Fahle, 2005) as well as of sequential dependencies in sentences (Gómez et al., 2006) and of grammatical categories (St. Clair & Monaghan, 2008). The effect of sleep on word learning has also been under recent investigation. Using priming experiments and fMRI, Gaskell and colleagues demonstrated that meanings of novel words can become relatively deeply integrated into the mental lexicon already after one night's sleep (but not after the same time delay excluding sleep): Words participated in lexical competition with familiar words and brain activity during this process was very similar to the brain activity elicited by familiar words (Dumay & Gaskell, 2007; Davis et al., 2008; Tamminen & Gaskell, 2008). While these studies give insight into the integration of novel words on a formal, phonological level, they do not examine consolidation on a more semantic and conceptual level. It is therefore a compelling question whether sleep has an effect on the generalization of novel words regarding their semantic categories.

Another finding concerning consolidation which may be relevant to the present research is that category associations can improve memory of items (Mandler, 1967; Gollin & Sharps, 1988). Given that learners who use SLCL may be sensitized to category membership, SLCL may also elicit better word consolidation, and therefore learning, than CSWL.

5.3. Experiment 5

5.3.1. Motivation

Experiment 5 set out to further investigate the interplay of CSWL and SLCL when information is independently applicable: That is, neither contrary as in Experiment 4, nor complementary as in Experiments 1 to 3 but redundantly co-present. The motivation for this manipulation was to complete the picture regarding the interaction of CSWL and SLCL and, importantly, to investigate the underlying nature of both mechanisms.

As in Experiment 4, we employed a learning setting in which each noun had two potential meanings, one with a co-occurrence frequency of 83% and one with a co-occurrence frequency of 50%. All distractors co-occurred only once (17%) with one noun. Additionally, we introduced the two-level factor Verb Type: Nouns were either in Condition *N(on-restrictive)* or in Condition *R(estrictive)*: In Condition N, nouns always followed non-restrictive verbs; in Condition R, nouns followed a restrictive verb in five of six trials. That means that, in Condition N, the only applicable word learning mechanism was CSWL whereas in Condition R, nouns could be learned using CSWL or SLCL (or both). Crucially, in Condition R, both CSWL and SLCL supported the 83% meaning (unlike in Experiment 4). Assuming that CSWL works in an incremental, parallel, and probabilistic manner, we hypothesized that participants would show sensitivity for both the difference between the co-occurrence of the 83% object and the 50% object and the difference between the 50% object and the 17% objects (distractors) in Condition N. Assuming that SLCL is more deterministic and direct than CSWL, we hypothesized that SLCL would block this parallel sensitivity in Condition R (when both CSWL and SLCL were independently applicable), supporting the 83% meaning. We moreover hypothesized that learners using SLCL (i.e. verb-based cues) would be sensitive to category associations because using the constraints of restrictive verbs is based on the knowledge about semantic categories. The most crucial differences of this design compared to the one of Experiment 4 were the following: In Condition R, both SLCL and CSWL supported the 83% referent and restrictive verbs were co-present with this object (recall paragraph *Noun Learning* of Section 4.1.2.2).

Additionally, Experiment 5 was designed to assess vocabulary retention after approximately 24 hours. Firstly, we investigated whether learners still know the learned nouns one day after learning (and how well). Secondly, we examined potential differences in the mental representation of a word meaning on Day 2 dependent on the mechanism that was used for learning. We hypothesized that learners may be more sensitive to semantic categories on Day 2 than on Day 1 due to enhanced generalization caused by consolidation (and sleep, e.g., St. Clair and Monaghan, 2008). We further hypothesized that this effect may be more pronounced when nouns were learned based on SLCL than when nouns were learned based on CSWL because SLCL may enhance sensitivity for category associations.

5.3.2. Methods

5.3.2.1. Participants

29 German native speakers, mainly students (from different disciplines), took part in Experiment 5 for a reimbursement of €10. Five of them had to be excluded due to verb learning problems. Data of the remaining 24 learners (5 males, 19 females, aged between 19 and 38, mean age 24) entered analyses. None of these participants had participated in either of Experiments 1 to 4.

5.3.2.2. Materials and Procedure

The experimental materials and procedure were similar to the ones in Experiment 4 (see Section 4.1.2.2). The miniature language consisted of 18 nouns (the two character names and ten object names of Experiment 4 plus six new object names), the same four verbs as in Experiment 4, and the same article as before.

On Day 1, the experiment consisted of the following phases: verb learning, verb testing, eye-tracker preparation and verb repetition (Phase 1), Noun-Learning Block 1 (Phase 2), Vocabulary Test 1 (Phase 3), Noun-Learning Block 2 (Phase 4), Vocabulary Test 2 (Phase 5), and a final verb check. The experiment lasted approximately 40 minutes. On Day 2, the vocabulary test of Phases 3 and 5 was repeated within one block (15 minutes).

Phase 1 was exactly as in Experiment 4: Participants were introduced to the experiment, verbs were trained and tested, and the eye-tracker was adjusted (see Paragraph *Verb Learning* of Section 4.1.2.2). Before learners were introduced into Phase 2, there was a short familiarization with the pictures of all objects before the sentence-phase started, in order to check for recognizability.

Noun Learning During noun learning (Phases 2 and 4), participants were exposed to pairs of static scenes and spoken SVO sentences (see Table 5.1). Novel nouns were embedded in the sentences as subjects and objects. Scenes contained animate characters and inanimate objects, partly including referents for the novel nouns. Instructions were very similar to the ones in Experiment 4 but less explicit: Participants were asked to understand the sentences and not explicitly told to learn the novel nouns. Noun learning consisted of 96 scene-sentence pairs, six presentations per object name. Eight object names were trained in Block 1 and eight in Block 2, which means that each block consisted of 48 trials. Each noun had two potential referents (i.e., meanings), one co-occurred with the noun in 83% of the trials (the high-frequency object, e.g., corn and hamburger in Table 5.1) and one co-occurred with the noun in 50% of the trials (the low-frequency object, e.g., socks and

Table 5.1.: Example trials Exp. 5

Verb Type	Cat.	Trial	Sentence	Scene
Non-restrictive	1	1	*Si laki gumbumema si daram* 'The man will point at the DARAM.'	corn (83%), socks (50%), gloves (17%), hat (17%), man
		2	*Si gadis tambamema si daram* 'The woman will take the DARAM.'	corn (83%), socks (50%), jumper (17%), scarf (17%), woman
		3	*Si gadis gumbumema si daram* 'The woman will point at the DARAM.'	corn (83%), socks (50%), top (17%), skirt (17%)
	2	4	*Si laki tambamema si daram* 'The man will take the DARAM.'	corn (83%), cap (17%), jacket (17%), dress (17%)
		5	*Si laki gumbumema si daram* 'The man will point at the DARAM.'	corn (83%), vest (17%), jeans (17%), coat (17%), man
	3	6	*Si gadis tambamema si daram* 'The woman will point at the DARAM.'	hamburger (17%), mushroom (17%), shorts (17%), apron (17%)
Restrictive	1	1	*Si laki bermamema si firel* 'The man will eat the FIREL.'	hamburger (83%), jumper (50%), gloves (17%), hat (17%), man
		2	*Si gadis bermamema si firel* 'The woman will eat the FIREL.'	hamburger (83%), jumper (50%), socks (17%), scarf (17%), woman
		3	*Si gadis bermamema si firel* 'The woman will eat the FIREL.'	hamburger (83%), jumper (50%), top (17%), skirt (17%)
	2	4	*Si laki bermamema si firel* 'The man will eat the FIREL.'	hamburger (83%), cap (17%), jacket (17%), dress (17%)
		5	*Si laki bermamema si firel* 'The man will eat the FIREL.'	hamburger (83%), vest (17%), jeans (17%), coat (17%), man
	3	6	*Si gadis tambamema si firel* 'The woman will point at the FIREL.'	corn (17%), mushroom (17%), shorts (17%), apron (17%)

jumper in Table 5.1). Objects other than the high-frequency object and the low-frequency object (i.e., distractors) all co-occurred only 17% with one noun. To avoid artifacts based on mutual exclusivity, no object belonged to the 83% group for one noun and to the 50% group for another noun (that is why there were twice as many objects as nouns). In addition to the different frequency groups, nouns were in one of two conditions: In Condition N, they always occurred with a non-restrictive verb. In Condition R, they occurred with a restrictive verb in 83% of their presentations (five of six). Importantly, in these restrictive trials, there was only one object depicted that matched the verbal restrictions. In contrast to Experiment 4, the SLCL and CSWL in Condition R supported the same referent: the 83% meaning (e.g., hamburger, Table 5.1). That means that, in Condition R, there were two cues for learning the meaning: co-occurrence frequencies and verb constraints.

The described manipulations resulted into three trial-type categories: In trials of Category 1 (three of six trials), both the high-frequency object and the low-frequency object were contained in the scene. In trials of Category 2 (two of six trials), the low-frequency object was not included in the scene and scenes belonging to trials of Category 3 (one of six trials) depicted neither the high-frequency nor the low-frequency referent. All scenes contained four objects. In scenes belonging to trials of Category 1 and Category 2, there was either one food object and three clothing objects or three food objects and one clothing object. In trials of Category 3, two food items and two clothing items were depicted. The combination of objects in scenes of Categories 1 and 2 (3 object of one category and 1 object of the other) was necessary to make sure that there was only one object matching the verbal constraints when the verb was restrictive. Category-3 scenes were arranged differently to make trials less predictive. As in Experiment 4, the character was not always included in the scene (but only in four of six trials) in order to implicitly remind participants of the fact that not everything that is mentioned in the sentence needs to be depicted.

Both Block 1 (Phase 2) and Block 2 (Phase 4) were subdivided into two parts for presentation. Each subpart consisted of 24 trials (four novel nouns). The order of presentation within these subparts was randomized. There was no vocabulary test or brake between the two subparts. We used this control in order to facilitate learning of the 18 new nouns. Pictures were counterbalanced in the same way as in the other experiments. Each of the 36 objects was presented 12 times, independent of condition or group, to avoid visual dominance effects.

Vocabulary Tests Importantly, the vocabulary tests in Phases 3 and 5 and on Day 2 were different than in the other experiments. Learners heard a noun and were asked to decide for one of four visual objects by clicking on it, as in Experiments 1 and 2. However, there were two different test types (see Table 5.2 for example trials). Either the 83% object, the 50% object, and two distractors were depicted (Test Type 1) or the 50% object and three

Table 5.2.: Example test trials Exp. 5

Verb Type	Test type	Noun	Objects
Non-restrictive	1	*si daram*	corn (83%), socks (50%), gloves (17%), scarf (17%)
	2	*si daram*	hamburger (CA), socks (50%), jumper (17%), apron (17%)
Restrictive	1	*si firel*	hamburger (83%), jumper (50%), gloves (17%), scarf (17%)
	2	*si firel*	apple (CA), jumper (50%), shorts (17%), apron (17%)

distractors were depicted, one of which matched the semantic category of the missing 83% object (from now on referred to as *category associate* or *CA*; Test Type 2). The forced choice was followed by a confidence rating as in Experiments 3 and 4, on a scale from 1 (not confident at all) to 9 (very confident). There were 24 test trials (12 per test type), each object name was used twice, once in each trial type, respectively.

Participants were presented training and testing trials according to one of four lists: Firstly, nouns were assigned to one of the two conditions. Secondly, objects were assigned to either the 83% group or the 50% group.

Day 2 Approximately 24 hours after this procedure (always including a night), the same learners returned to the lab. The eye-tracker was adjusted and the vocabulary test was repeated, with the same trials as on Day 1, presented in random order in one block. Participants were not informed about the reason for the Day-2 session on Day 1, neither were they requested to practice or memorize what they had learned on Day 1 in any way.

5.3.3. Predictions

For decisions in vocabulary-test trials of Test Type 1, we made the following predictions: Hypothesizing that learners in Condition R are guided by SLCL and learners in Condition N are guided by CSWL and given that both CSWL and SLCL supported the 83% meaning, we predicted that learners would choose the 83% object (e.g., corn) more often than the other objects in both conditions. However, hypothesizing that SLCL is more deterministic than CSWL, we expected a clearer dominance for the 83% meaning in Condition R than in Condition N. Confidence ratings were expected to reflect this trend: We predicted higher ratings in Condition R than in Condition N. For N, we predicted that a secondary preference for the 50% meaning (socks) may be reflected in decisions: Assuming that CSWL operates in parallel, learners should be sensitive to the difference in co-occurrence frequency between the 50% object and the 17% distractors. However, we expected that this secondary preference may potentially not become visible due to the presence of the 83% object.

In vocabulary-test trials of Test Type 2, on the other hand, we predicted a tendency for learners to choose the 50% meaning (socks) in Condition N, hypothesizing that CSWL works parallel: Based on co-occurrence frequencies, it was the most plausible alternative to the non-present 83% meaning. For R, on the contrary, we expected learners not to be sensitive for the difference between co-occurrence frequencies of 50% and 17%. Instead we predicted learners to prefer the category associate (hamburger), hypothesizing that learning via SLCL enhances sensitivity for category membership.

We assumed that test trials of Type 2 are more challenging than test trials of Type 1 because learners potentially encounter falsification of their hypotheses about word meanings.

We therefore predicted confidence ratings to be higher in Type-1 decisions than in Type-2 decisions.

Eye-movements during learning (Trial Category 1) and testing (Test Type 1) were predicted to support the pattern of off-line results: For Condition R, we expected a clear preference for inspecting the 83% object whereas, for Condition N, a secondary preference for inspecting the 50% object was predicted. Given the hypothesized determinism of SLCL, operating in Condition R, and the hypothesized probabilistic nature of CSWL, operating in Condition N, we additionally expected a more pronounced preference to look at the 83% object during learning trials of Type 2 in Condition R than Condition N. Finally, in test trials of Type 2, we predicted a preference in looks for the category associate in Condition R and a bias to inspect the 50% candidate in Condition N. Based on our findings from Experiment 1 to 4, we also expected anticipatory eye-movements during the verb region in Condition R in learning trials of both Trial Categories 1 and 2.

We further predicted that learning as evidenced by decisions for the 83% meaning in Test Type 1 would still be above chance in both conditions on Day 2, hypothesizing that SLCL and CSWL are both learning mechanisms which can result in deep learning. However, we predicted performance in the vocabulary test to be worse on Day 2 than on Day 1, due to general memory decay (*forgetting curve*, Baddeley, 1998). We expected this decline in performance to be higher in Condition N than Condition R. This prediction was due to our hypothesis that choosing the low-frequency candidate would be an alternative in Condition N but not in Condition R: Given that memory decays, the sensitivity for the difference between the 83% candidate and the 50% may be blurred on Day 2.

Hypothesizing that over-night consolidation improves generalization, we moreover predicted learners to choose the category associate in Test Type 2 more often on Day 2 than on Day 1. This increase was expected to be higher for Condition R than Condition N because learning in Condition R was hypothesized to rely on SLCL which, again, was hypothesized to enhance sensitivity for category associations. Finally, confidence ratings were expected to be lower on Day 2 than Day 1, as well, due to people's assumed awareness of memory decay.

5.3.4. Data Analysis and Results

5.3.4.1. Off-line Results

Learning Rates Learning rates (83%-candidate chosen in Test Type 1) were significantly above chance (25%) for both verb types and on both days (see Table 5.3; Day 1, N: $t(23) = 7.995, p < .001$; Day 1, R: $t(23) = 16.284, p < .001$; Day 2, N: $t(23) = 4.978, p < .001$; Day 2, R: $t(23) = 10.791, p < .001$).

Differences in learning rates (Test Type 1) and chosen meanings (Test Types 1 and 2)

Table 5.3.: Learning rates Exp. 5

	Non-restrictive	Restrictive	Both conditions
Day 1	60%	84%	72%
Day 2	49%	78%	63%
Both	55 %	80%	68%

between verb types and days were analyzed with linear mixed-effects models, using logistic regression for categorical data (logit link function) and linear regression for continuous data, with participant and item as random factors (see Section 3.1.4.2 for explanations about this method).

Learning was significantly better on Day 1 than Day 2 ($\chi(1) = 7.33, p < .01$; Table 5.5, Row 2) and in Condition R than Condition N ($\chi(1) = 68.671, p < .001$; Table 5.5, Row 5). Although the interaction between factors Day and Verb Type was not significant ($\chi(1) = 0.136, p = .712$), we found that the decrease in learning rate was reliable in Condition N ($\chi(1) = 6.078, p < .050$; Table 5.4, Rows 1-2) but did not reach significance in Condition R ($\chi(1) = 2.562, p = .109$; Table 5.4, Rows 3-4).

For analyzing differences in the frequencies of decisions for the four depicted objects in the vocabulary test (and how these differences differed in verb types), we conducted two-way repeated measures ANOVAS with VerbType (N/R) and Chosen Meaning (Test Type 1: 83%, 50%, Distractor 1, Distractor 2; Test Type 2: CA, 50%, Distractor 1, Distractor 2) averaged across items (F_1) and subjects (F_2). The reason for not using linear mixed effect models was that multilevel logistic regression, which would have been required due to the fact that Chosen Meaning has four levels, was not implemented for the analysis software we were using (statistical package R).

Chosen Meanings in Test Type 1 First, data from test trials of Test Type 1 (83% candidate and 50% candidate present) was analyzed (see Figure 5.1). ANOVAS with Verb Type and Chosen Meaning as fixed factors revealed significant main effects of Chosen Meaning for both days (Day 1: $F_1(3, 69) = 177.800, p < .001; F_2(3, 45) = 304.103, p < .001$; Day 2: $F_1(3, 69) = 63.008, p < .001; F_2(3, 45) = 122.162, p < .001$): The 83% object was chosen significantly more often than the other objects. Verb Type and Chosen Meaning also interacted (Day 1: $F_1(3, 69) = 12.293, p < .001; F_2(3, 45) = 12.563, p < .001$; Day 2: $F_1(3, 69) = 20.906, p < .001; F_2(3, 45) = 24.011, p < .001$). Tests of Within-Subjects Contrasts indicate that the interactions on both days were caused by a greater

Table 5.4.: Lmer models for learning rates (83%-candidate choices in Test Type 1) for single conditions on Days 1 and 2, Exp. 5
$83\%chosen \sim day + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
Condition N					
1	(Int) (*Day1*)	0.506	0.262	1.930	$= .054$
2	*Day2*	−0.556	0.219	−2.53	$< .050$
Condition R					
3	(Int) (*Day1*)	2.111	0.369	5.716	$< .001$
4	(Int) (*Day2*)	−0.468	0.288	−1.629	$= .103$

Table 5.5.: Lmer models for learning rates (83%-candidate choices in Test Type 1), Exp. 5
m1: $83\%chosen \sim day + (1|sub) + (1|item), family = binomial(link = "logit")$
m2: $83\%chosen \sim day + condition + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
m1					
1	(Int) (*Day1*)	1.063	0.209	5.081	$< .001$
2	*Day2*	−0.446	0.163	−2.738	$< .001$
m2					
3	(Int) (*Day1-N*)	0.466	0.240	1.942	$= .052$
4	(Int) (*Day2*)	−0.495	0.172	−2.877	$< .010$
5	*R*	1.438	0.178	8.068	$< .001$

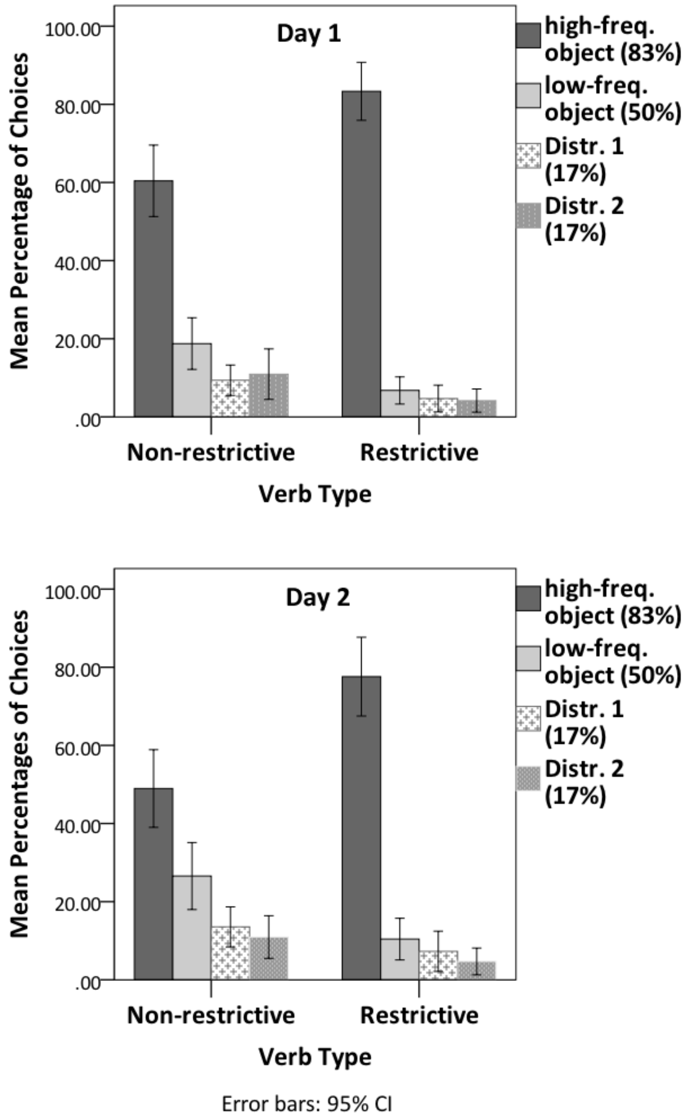


Figure 5.1.: Chosen meanings in vocabulary-test trials of Test Type 1 on Day 1 (top chart) and Day 2 (bottom chart), Exp. 5

difference in R than N between the frequencies of 83% selections and 50% selections (Day 1: $F_1(1,23) = 16.422, p < .001; F_2(1,15) = 15.007, p < .001$; Day 2: $F_1(1,23) = 28.145, p < .001; F_2(1,15) = 44.518, p < .001$).

To further evaluate these interactions between Verb Type and Chosen Meaning, we conducted one-way ANOVA analyses with Chosen Meaning as fixed factor, separately for conditions and days. Main effects were found for both conditions and on both days (Day

1, N: $F_1(3, 69) = 40.272, p < .001; F_2(3, 45) = 51.277, p < .001$; Day 1, R: $F1(3, 69) = 225.448, p < .001; F_2(3, 45) = 304.856, p < .001$; Day 2, N: $F1(3, 69) = 16.933, p < .001; F_2(3, 45) = 26.215, p < .001$; Day 2, R: $F1(3, 69) = 94.362, p < .001; F_2(3, 45) = 163.944, p < .001$). Analyses revealed that the 83% object was chosen significantly more often than each other object in both conditions and on both days (see Table 5.6; Day 1, N: Rows 1-3; Day 1, R: Rows 13-15; Day 2, N: Rows 25-27; Day 2, R: Rows 37-39). We additionally found for Condition N on Day 2 that learners chose the low-frequency object (marginally) significantly more often than the distractor objects (Table 5.6, Rows 29-30).

We also evaluated the interactions between Verb Type and Chosen Meaning by entering the data of only 83% choices and 50% choices into linear mixed effect models, using logistic regression. We found VerbType to be a predictor for Chosen Meaning (low frequency or high frequency) on both days (Day 1: $\chi(1) = 17.300, p < .001$; Day 2: $\chi(1) = 25.739, p < .001$): The high-frequency object was chosen reliably more often in Condition R than in Condition N (Day 1: $\chi(1) = 28.420, p < .001$, Table 5.7, Row 6; Day 2: $\chi(1) = 41.449, p < .001$, Table 5.7, Row 8) and the low-frequency object was picked reliably more often in Condition N than in Condition R (Day 1: $\chi(1) = 12.688, p < .001$, Table 5.7, Row 10; Day 2: $\chi(1) = 17.940, p < .001$ Table 5.7, Row 12).

Finally, to directly compare the effects of the factors Day and Chosen Meaning, we conducted two-level repeated measures ANOVAs, separately for Conditions N and R. We found a significant interaction between Day and Chosen Meaning for Condition N ($F_1(3, 69) = 5.345, p < .010; F_2(3, 45) = 8.578, p < .001$): While participants chose the high-frequency object more often on Day 1 than Day 2, the low-frequency candidate was chosen more often on Day 2 than on Day 1. When entering only high-frequency object choices and low-frequency object choices as dependent variable into a linear mixed effect model with logit link function, we also found that Day predicted which object was chosen ($\chi(1) = 4.802, p < .050$; see Table 5.8). No interaction was found for Condition R ($F_1(3, 69) = 1.959, p = .128; F_2(3, 45) = 2.427, p < .078$).

Chosen Meanings in Test Type 2 For Test Type 2 (50%-object and category associate available; Figure 5.2), we found a main effect of Chosen Meaning (Day 1: $F_1(3, 69) = 15.930, p < .001; F_2(3, 45) = 19.722, p < .001$; Day 2: $F_1(3, 69) = 14.495, p < .001; F_2(3, 45) = 20.250, p < .001$) and an interaction between factors Verb Type and Chosen Meaning for both days (Day 1: $F_1(3, 69) = 6.762, p < .001; F_2(3, 45) = 7.007, p < .010$; Day 2: $F_1(3, 69) = 14.800, p < .001; F_2(3, 45) = 15.652, p < .001$). Tests of Within-Subjects Contrasts indicate that the interactions were partly caused by differences in Verb Type concerning the frequencies of CA choices versus 50% choices.

To evaluate the interactions, we again conducted one-way ANOVAS with Chosen Meaning as fixed factor, separately for conditions and days and found that Chosen Meaning had a

Table 5.6.: Pairwise comparisons for ANOVAs by subject (Bonferroni adjustment), Test Type 1, Exp. 5

	chosen meaning	chosen meaning	Mean Difference	SE	*p*
Day 1, Condition N					
1	83%	50%	.417	.067	< .001
2	83%	17%-1	.510	.053	< .001
3	83%	17%-2	.495	.070	< .001
4	50%	83%	-.417	.067	< .001
5	50%	17%-1	.094	.042	= .215
6	50%	17%-2	.078	.046	= .603
7	17%-1	83%	-.510	.053	< .001
8	17%-1	50%	-.094	.042	= .215
9	17%-1	17%-2	.016	.035	= 1.000
10	17%-2	83%	-.495	.070	< .001
11	17%-2	50%	-.078	.046	= .603
12	17%-2	17%-1	.016	.035	= 1.000
Day 1, Condition R					
13	83%	50%	.766	.050	< .001
14	83%	17%-1	.786	.050	< .001
15	83%	17%-2	.792	.047	< .001
16	50%	83%	-.766	.050	< .001
17	50%	17%-1	.021	.018	= 1.000
18	50%	17%-2	.026	.020	= 1.000
19	17%-1	83%	-.786	.050	< .001
20	17%-1	50%	-.021	.018	= 1.000
21	17%-1	17%-2	.005	.016	= 1.000
22	17%-2	83%	-.792	.047	< .001
23	17%-2	50%	-.026	.020	= 1.000
24	17%-2	17%-1	-.005	.016	= 1.000
Day 2, Condition N					
25	83%	50%	.224	.084	= .085
26	83%	17%-1	.354	.063	< .001
27	83%	17%-2	.380	.059	< .001
28	50%	83%	-.224	.084	= .085
29	50%	17%-1	.130	.047	= .062
30	50%	17%-2	.156	.054	< .050
31	17%-1	83%	-.354	.063	< .001
32	17%-1	50%	-.130	.047	= .062
33	17%-1	17%-2	.026	.041	= 1.000
34	17%-2	83%	-.380	.059	< .001
35	17%-2	50%	-.156	.054	< .050
36	17%-2	17%-1	.026	.041	= 1.000
Day 2, Condition R					
37	83%	50%	.672	.072	< .001
38	83%	17%-1	.703	.070	< .001
39	83%	17%-2	.729	.058	< .001
40	50%	83%	-.672	.072	< .001
41	50%	17%-1	.031	.025	= 1.000
42	50%	17%-2	.057	.028	= .320
43	17%-1	83%	-.703	.070	< .001
44	17%-1	50%	-.031	.025	= 1.000
45	17%-1	17%-2	.026	.029	= 1.000
46	17%-2	83%	-.729	.058	< .001
47	17%-2	50%	-.057	.028	= .320
48	17%-2	17%-1	-.026	.029	= 1.000

Table 5.7.: Lmer models for chosen meanings in conditions (low-frequency vs. high-frequency object, high-frequency meaning choices, low-frequency meaning choices), TestType 1, Exp. 5
$chosen \sim VerbType + (1|sub) + (1|item), family = binomial(link = "logit")$

		Predictor	Coef.	SE	Wald z	p
low vs. high frequency						
1	*Day 1*	(Int) (N)	−1.170	0.191	−6.133	$< .001$
2		R	−1.340	0.346	−3.876	$< .001$
3	*Day 2*	(Int) (N)	−0.690	0.268	−2.576	$< .010$
4		R	−1.534	0.313	−4.901	$< .001$
amount high-frequency choices						
5	*Day 1*	(Int) (N)	0.465	0.188	2.474	$< .050$
6		R	1.312	0.254	5.160	$< .001$
7	*Day 2*	(Int) (N)	−0.052	0.277	0.186	$= .852$
8		R	−1.577	0.248	−6.347	$< .001$
amount low-frequency choices						
9	*Day 1*	(Int) (N)	−1.460	0.185	−7.891	$< .001$
10		R	−1.151	0.342	−3.369	$< .001$
11	*Day 2*	(Int) (N)	−1.105	0.210	−5.250	$< .001$
12		R	−1.200	0.300	−4.048	$< .001$

Table 5.8.: Lmer models for chosen meanings (low-frequency vs. high-frequency object) on days, TestType 1, Condition N, Exp. 5
$chosen \sim Day + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
1	(Int) (*Day 1*)	−1.338	0.292	−4.582	$< .001$
2	*Day 2*	0.614	0.275	2.234	$< .050$

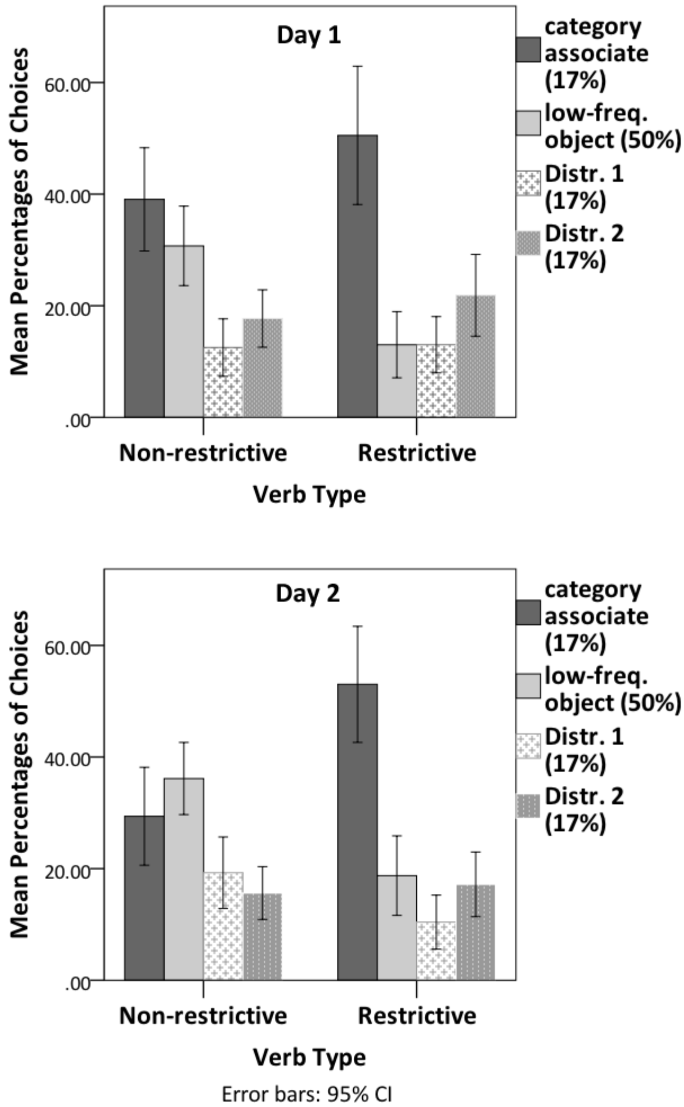


Figure 5.2.: Chosen meanings in trials of Test Type 2 on Day 1 (top chart) and Day 2 (bottom chart), Exp. 5

significant effect for both verb types and on both days (Day 1, N: $F_1(3,69) = 9.938, p < .001; F_2(3,45) = 9.018, p < .001$; Day 1, R: $F1(3,69) = 15.165, p < .001; F_2(3,45) = 22.132, p < .001$; Day 2, N: $F1(3,69) = 6.218, p < .010; F_2(3,45) = 5.686, p < .010$; Day 2, R: $F1(3,69) = 21.573, p < .001; F_2(3,45) = 39.531, p < .001$). Pairwise comparisons reveal notable differences between verb types and days. In Condition R, the category associate was chosen significantly more often than each other object on both days (Table 5.9, Day 1: Rows 13-15, Day 2: Rows 37-39); no other difference was significant (Table 5.9, Day 1: Rows 16-24, Day 2: Rows 40-48). In Condition N, in contrast, the category associate was chosen significantly more often than the two distractor objects on Day 1 but not on Day 2 (Table 5.9, Day 1: Rows 2-3, Day 2: Rows 26-27). The 50% object, on the contrary, was selected significantly more often than the two distractor objects on both days (Table 5.9, Day 1: Rows 5-6, Day 2: Rows 29-30). The difference between CA choices and 50%-object choices was not significant on either day (Table 5.9, Day 1: Row 1, Day 2: Row 25).

Again, interactions between Verb Type and Chosen Meaning were additionally explored by entering the data of only CA choices and 50% choices into linear mixed effect models, using logistic regression. We found that Verb Type predicts whether the low-frequency object or the category associate is chosen (Day 1: $\chi(1) = 15.651, p < .001$, Table 5.10, Row 2; Day 2: $\chi(1) = 26.869, p < .001$, Table 5.10 , Row 4): On both days, learners made significantly more category-competitor decisions in Condition R than in Condition N (Day 1: $\chi(1) = 7.010, p < .01$, Table 5.10, Row 6; Day 2: $\chi(1) = 24.402, p < .001$, Table 5.10 , Row 8) and significantly more low-frequency choices in Condition N than in Condition R (Day 1: $\chi(1) = 17.612, p < .001$, Table 5.10, Row 10; Day 2: $\chi(1) = 15.007, p < .001$; Table 5.10, Row 12).

We then directly evaluated the effect of Day on decisions for the two verb types by conducting two-level repeated measures ANOVAs with factors Day and Chosen Meaning, separately for Conditions N and R. We found a (marginally) significant interaction for Condition N ($F_1(3,69) = 4.044, p < .050; F_2(3,45) = 2.570, p = .066$) but no such effect for Condition R ($F_1(3,69) = 1.186, p = .321; F_2(3,45) = 9.018, p < .343$). Finally, based on our finding that the preference for the category associate was different between days in Condition R, we tested whether the number of category-associate choices in the two verb types was stable across days. CA choices were entered into linear mixed effect models (using logistic regression) as predictors and Verb Type and Day as fixed factors. We found a marginally significant interaction ($\chi(2) = 4.998, p = .082$): While the number of CA choices was stable across days in Condition R ($\chi(1) = 0.205, p = .650$, Table 5.11, Rows 3-4), there was a significant decrease from Day 1 to Day 2 in Condition N ($\chi(1) = 4.930, p < .05$; Table 5.11, Rows 1-2).

Table 5.9.: Pairwise comparisons for ANOVAs by subject (Bonferroni adjustment), Test Type 2, Exp. 5

	chosen	chosen	Mean Difference	SE	*p*
Day 1, Condition N					
1	CA	50%	.083	.074	= 1.00
2	CA	17%-1	.266	.058	< .010
3	CA	17%-2	.214	.060	< .050
4	50%	83%	−.083	.074	= 1.00
5	50%	17%-1	.182	.045	< .010
6	50%	17%-2	.130	.042	< .050
7	17%-1	CA%	-.266	.058	< .010
8	17%-1	50%	-.182	.045	< .010
9	17%-1	17%-2	-.052	.039	= 1.000
10	17%-2	CA%	-.214	.060	< .010
11	17%-2	50%	-.130	.042	< .050
12	17%-2	17%-1	.052	.039	= 1.000
Day 1, Condition R					
13	CA	50%	.375	.079	< .010
14	CA	17%-1	.375	.079	< .010
15	CA	17%-2	.286	.089	< .050
16	50%	83%	−.375	.079	< .010
17	50%	17%-1	.000	.032	= 1.00
18	50%	17%-2	−.089	.046	= .402
19	17%-1	CA%	-.375	.079	< .001
20	17%-1	50%	.000	.032	= .215
21	17%-1	17%-2	-.089	.037	= 1.000
22	17%-2	CA%	-.286	.089	< .050
23	17%-2	50%	.089	.046	= .402
24	17%-2	17%-1	.089	.037	= .156
Day 2, Condition N					
25	CA	50%	−.068	.067	= 1.00
26	CA	17%-1	.099	.064	= .808
27	CA	17%-2	0.135	.050	= .079
28	50%	83%	.068	.067	= 1.00
29	50%	17%-1	.167	.042	< .010
30	50%	17%-2	0.203	.040	< .001
31	17%-1	CA%	-.099	.064	= .808
32	17%-1	50%	-.167	.042	< .010
33	17%-1	17%-2	.036	.046	= 1.000
34	17%-2	CA%	-.135	.050	= .079
35	17%-2	50%	-.203	.040	< .001
36	17%-2	17%-1	-.036	.046	= 1.000
Day 2, Condition R					
37	CA	50%	.339	.078	< .010
38	CA	17%-1	.422	.066	< .001
39	CA	17%-2	.354	.066	< .001
40	50%	83%	−.339	.078	< .010
41	50%	17%-1	.083	.038	= .235
42	50%	17%-2	.016	.049	= 1.00
43	17%-1	CA%	-.422	.066	< .001
44	17%-1	50%	-.083	.038	= .235
45	17%-1	17%-2	-.068	.035	= .404
46	17%-2	CA%	-.354	.066	< .001
47	17%-2	50%	-.016	.049	= 1.000
48	17%-2	17%-1	.068	.035	= .404

Table 5.10.: Lmer models for chosen meanings in conditions (low-frequency object/category associate, high-frequency meaning choices, low-frequency meaning choices), Test Type 2, Exp. 5
$chosen \sim VerbType + (1|sub) + (1|item), family = binomial(link = "logit")$

		Predictor	Coef.	SE	Wald z	p
low vs. high frequency						
1	*Day 1*	(Int) (N)	−0.224	0.255	−0.877	= .380
2		R	−1.177	0.301	−3.917	< .001
3	*Day 2*	(Int) (N)	0.3009	0.282	1.067	= .286
4		R	−1.486	0.284	−5.234	< .001
amount cat-as. choices						
5	*Day 1*	(Int) (N)	−0.543	0.264	−2.058	< .050
6		R	0.605	0.223	2.711	< .010
7	*Day 2*	(Int) (N)	−0.992	0.224	−4.434	< .001
8		R	1.107	0.224	4.948	< .001
amount low-frequency choices						
9	*Day 1*	(Int) (N)	−0.830	0.169	−4.923	< .001
10		R	−1.084	0.267	−4.050	< .001
11	*Day 2*	(Int) (N)	−0.609	0.187	−3.255	< .010
12		R	−0.926	0.242	−3.833	< .001

Table 5.11.: Lmer models for number of CA-Choices on Days 1 and 2 for single conditions, Trial Type 2, Exp. 5
$83\% chosen \sim day + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
Condition N					
1	(Int) (*Day1*)	−0.536	0.262	−2.045	< .050
2	*Day2*	−0.522	0.229	−2.275	< .050
Condition R					
3	(Int) (*Day1*)	0.055	0.264	0.209	= .834
4	(Int) (*Day2*)	0.103	0.222	0.464	= .643

Table 5.12.: Lmer models & p-values from MCMC sampling for confidence ratings, differences test types and days, Exp. 5
$m1 : confidencerating \sim testtype + (1|sub) + (1|item)$
$m2 : confidencerating \sim day + (1|sub) + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
m1							
1	(Int.) (Type1)	6.019	0.224	26.91	6.020	$< .001$	$< .001$
2	Type 2	-2.764	0.202	-13.66	-2.765	$< .001$	$< .001$
m2							
5	(Int.) (Day1)	4.751	0.325	14.610	4.749	$< .001$	$< .001$
6	Day2	-0.225	0.125	-1.808	-0.224	0.077	0.071

Confidence Ratings Confidence ratings were analyzed as in Experiments 3 and 4 (see Section 3.3.4.1). Learners rated their confidence as significantly higher for Test Type 1 than Test Type 2 ($\chi(1) = 62.184, p < .001$; Table 5.12, Rows 1-2) and as marginally significantly higher on Day 1 than Day 2 ($\chi(1) = 3.267, p = .071$; Table 5.12, Rows 3-4). See Figures 5.3 and 5.4.

Additionally, we found for both days that, for Trial Type 1, ratings were higher in Condition R than in Condition N (Day 1: $\chi(1) = 31.008, p < .001$, Table 5.13, Row 2; Day 2: $\chi(1) = 54.940, p < .001$, Table 5.13, Row 6) and higher for the high-frequency object choices than for the low-frequency object choices (Day 1: $\chi(1) = 63.138, p < .001$, Table 5.13, Row 4; Day 2: $\chi(1) = 39.944, p < .001$, Table 5.13, Row 8).

For Trial Type 2, on the other hand, ratings were higher for Condition N than for Condition R (Day 1: $\chi(1) = 15.030, p < .001$, Table 5.13, Row 10; Day 2: $\chi(1) = 6.042, p < .05$, Table 5.13, Row 14) and also higher for low-frequency object choices than the category associate choices (Day 1: $\chi(1) = 4.440, p < .05$ Table 5.13, Row 12; Day 2: $\chi(1) = 13.488, p < .001$, Table 5.13, Row 16). Confidence ratings for Trial Type 2 for single days and single conditions are summarized in Figure 5.4. Interestingly, for Day 1, factors Verb Type and Chosen Meaning interact marginally ($\chi(1) = 3.202, p = .074$): The difference between ratings for the low-frequency object and the category associate is significantly larger in Condition R than in Condition N (and ratings for the low-frequency competitor are higher than ratings for the category associate in Condition N but lower in Condition R).

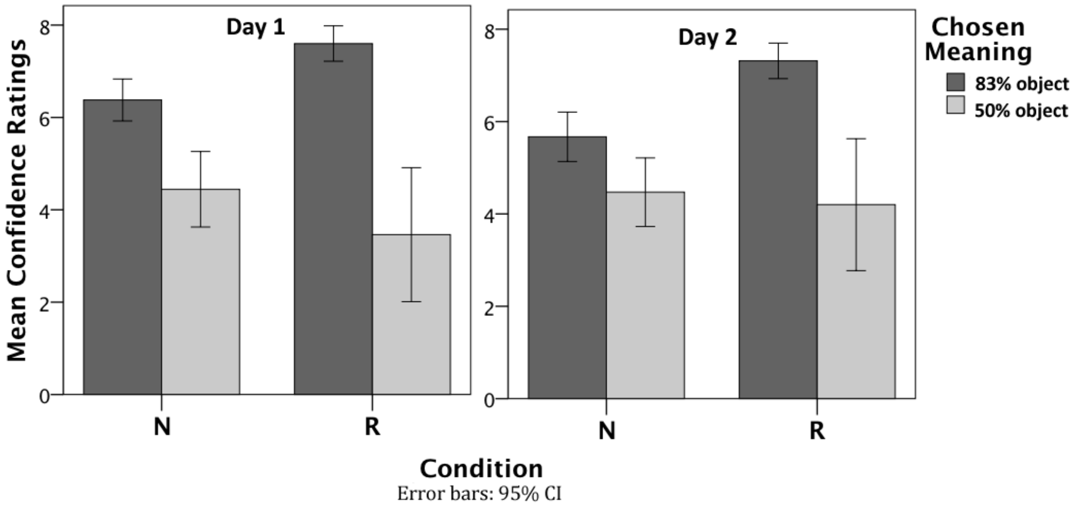


Figure 5.3.: Confidence ratings in trials of Test Type 1 on Day 1 (left chart) and Day 2 (right chart), Exp. 5

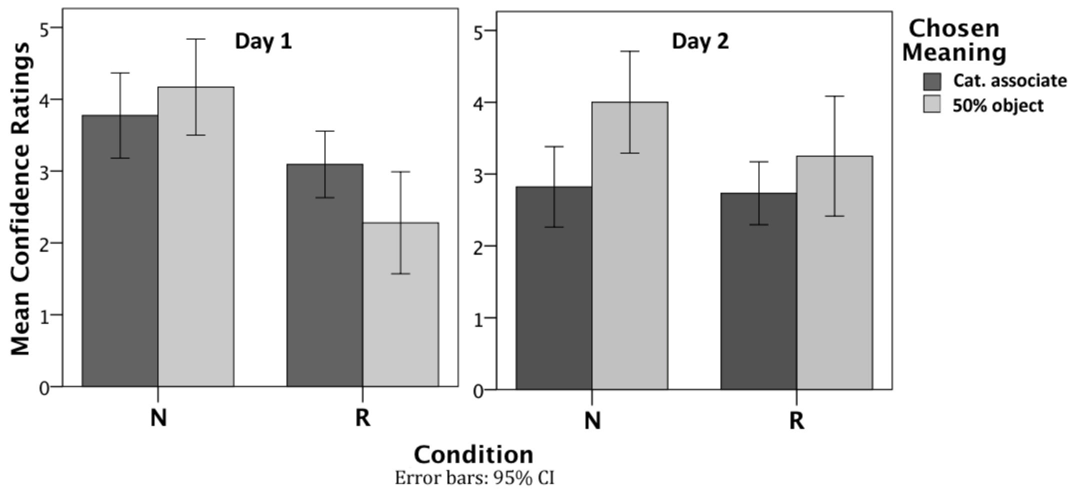


Figure 5.4.: Confidence ratings in trials of Test Type 2 on Day 1 (left chart) and Day 2 (right chart), Exp. 5

Table 5.13.: Lmer models & p-values from MCMC sampling for confidence ratings, Exp. 5
$VerbType \sim confidencerating + (1|sub) + (1|item)$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
Test Type 1, Day 1							
1	(Int.) (N)	5.516	0.296	18.620	5.518	< .001	< .001
2	R	1.461	0.257	5.69	1.460	< .001	< .001
3	(Int.) (83%)	7.069	0.292	24.251	7.074	< .001	< .001
4	50%	−2.977	0.356	−8.375	−2.984	< .001	< .001
Test Type 1, Day 2							
5	(Int.) (N)	4.901	0.277	17.705	4.901	< .001	< .001
6	R	1.802	0.257	7.003	1.802	< .001	< .001
7	(Int.) (83%)	6.679	0.206	32.480	6.675	< .001	< .001
8	50%	−2.336	0.358	−6.520	−2.316	< .001	< .001
Test Type 2, Day 1							
9	(Int.) (N)	3.630	0.309	11.763	3.63.	< .001	< .001
10	R	−0.752	0.192	−3.915	−0.750	< .001	< .001
11	(Int.) (50%)	3.273	0.339	9.645	3.290	< .001	< .001
12	cat.as.	0.627	0.293	2.139	0.580	= .055	< .050
Test Type 2, Day 2							
13	(Int.) (N)	3.490	0.345	10.117	3.488	< .001	< .001
14	R	−0.479	0.194	−2.465	−0.481	< .050	< .050
15	(Int.) (50%)	2.856	0.368	7.750	2.844	< .001	< .001
16	cat.as.	0.985	0.264	3.725	0.982	< .001	< .001

5.3.4.2. On-line Results

Eye movements during noun learning were analyzed for regions verb and NP2, separately for Trial Categories 1 and 2 and for both conditions. Data was treated as in Experiment 4 (see Section 4.1.4.2).

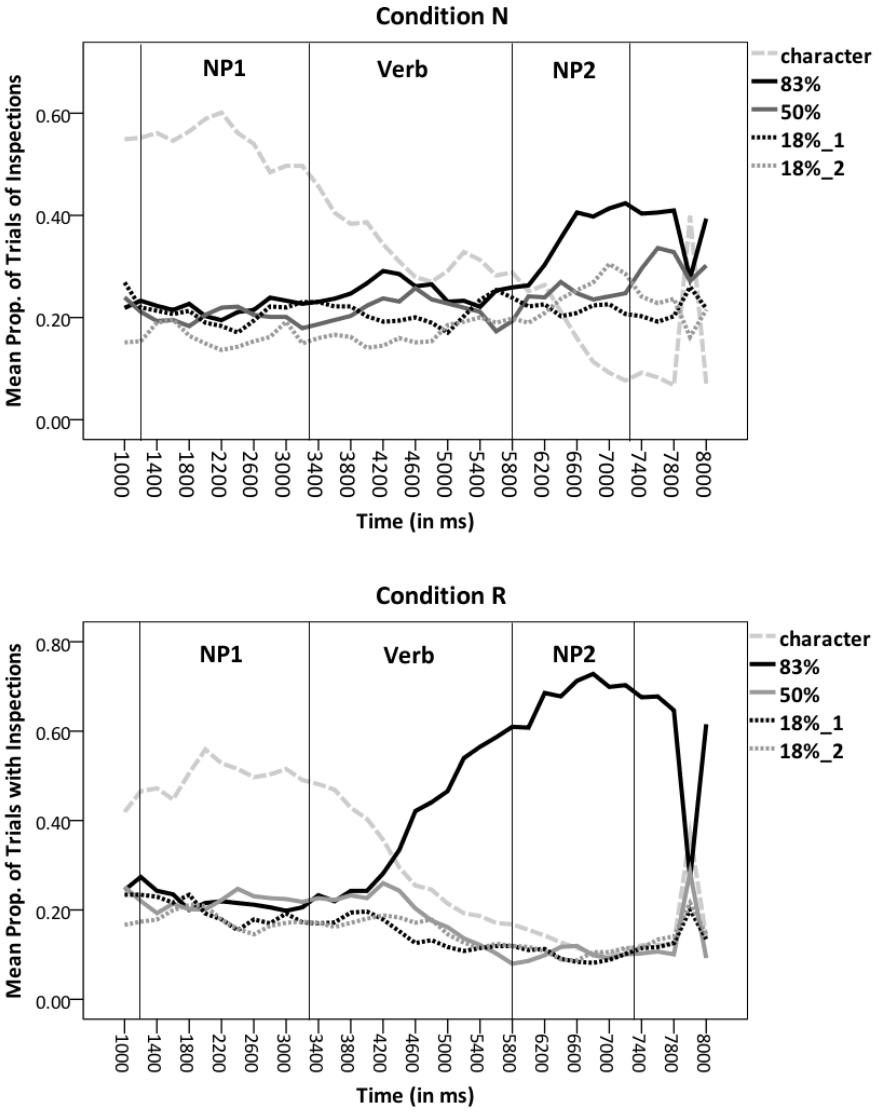


Figure 5.5.: Timegraphs Exp. 5, Trial Category 1, Condition N (top) and Condition R (bottom)

Eye-movements during Verb and NP2 in Trials of Trial Category 1 For Trial Category 1 (high-frequency object and low-frequency object included in the scene), we found effects of ROI in verb region of Condition R ($\chi(3) = 108.700, p < .001$, Table 5.14, Rows 10-12) but not in verb region of Condition N ($\chi(3) = 6.258, p = .100$, Table 5.14, Rows

2-4) and in NP2 of both conditions (N: $\chi(3) = 13.463, p < .01$, Table 5.14 , Rows 6-8; R: $\chi(3) = 97.157, p < .001$, Table 5.14, Rows 14-16): The high-frequency object was inspected reliably more often than any other object (i.e., the low-frequency object and both distractors). There was an interaction between ROI and Verb Type for both verb region ($\chi(4) = 40.48, p < .001$) and NP2 ($\chi(3) = 28.366, p < .001$; Figure 5.6). Crucially, this interaction is still present when only looks to the high-frequency object and looks to the low-frequency object are included (verb: $\chi(1) = 30.529, p < .001$; NP2: $\chi(1) = 11.026, p < .001$): The difference between looks to both objects was reliably larger in Condition R than Condition N. A further analysis revealed that learners made significantly more looks to the low-frequency object during NP2 in Condition N than in Condition R ($\chi(1) = 62.615, p < .001$, Table 5.15). See Figure 5.5 for timegraphs.

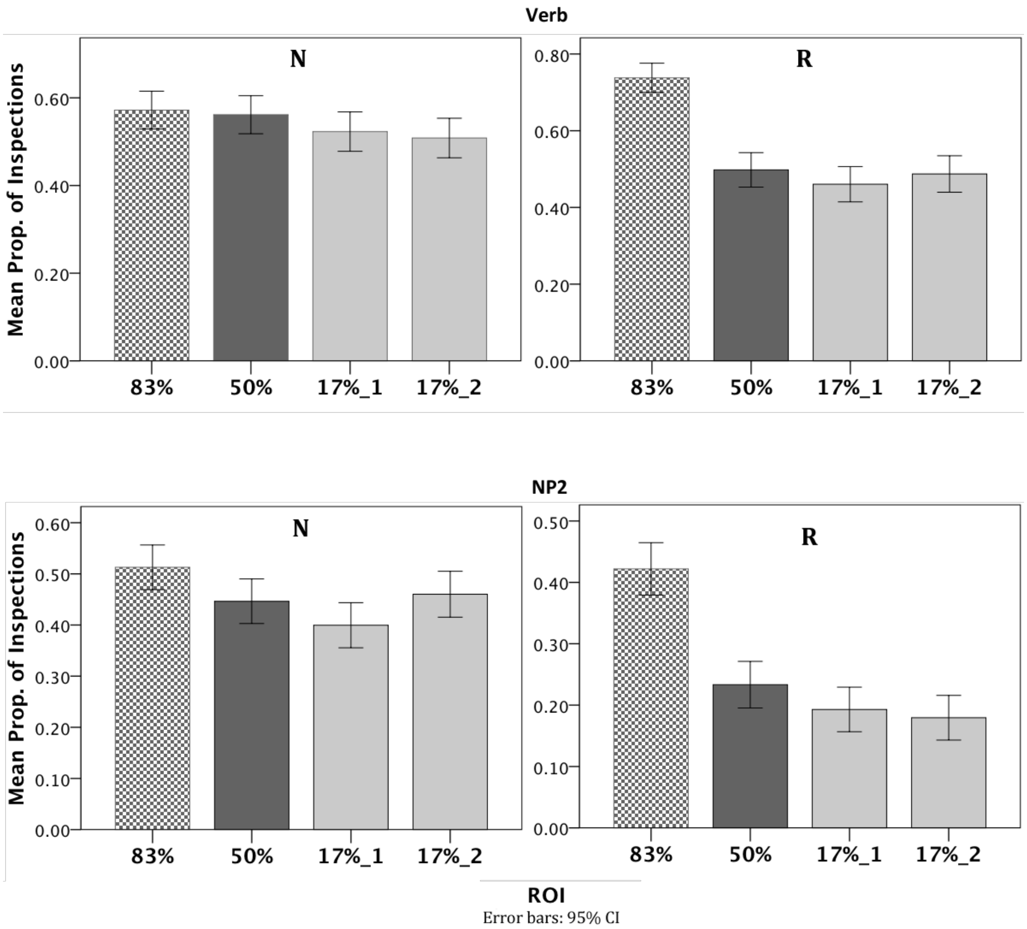


Figure 5.6.: Proportions of trials with at least one inspections to ROIs during the verb (top charts) and NP2 (bottom charts) in trials of Category 1, Day 1, Exp. 5: Condition N (left charts) and Condition R (right charts)

Table 5.14.: Lmer models for inspections on ROIs during the verb and NP2, Trial Category 1, Exp. 5
$Inspections \sim ROI + (1|sub) + (1|item) + (1|block), family = binomial(link = "logit")$

		Predictor	Coef.	SE	Wald z	p
Condition N						
1	*V*	(Int) (high-f.)	0.298	0.125	2.378	$< .050$
2		low-f.	−0.056	0.129	−0.436	$= .663$
3		17%-1	−0.223	0.130	−1.711	$= .087$
4		17%-1	−0.283	0.130	−2.168	$< .050$
5	*NP2*	(Int) (high-f.)	0.050	0.105	0.481	$= .630$
6		low-f.	−0.276	0.127	−2.176	$< .050$
7		17%-1	−0.472	0.130	−3.640	$< .001$
8		17%-1	−0.225	0.129	−1.752	$= .080$
Condition R						
9	*V*	(Int) (high-f.)	1.066	0.123	8.639	$< .001$
10		low-f.	1.086	0.137	−7.937	$< .001$
11		17%-1	−1.233	0.139	−8.880	$< .001$
12		17%-1	−1.138	0.141	−8.085	$< .001$
13	*NP2*	(Int) (high-f.)	−0.336	0.124	−2.702	$< .010$
14		low-f.	−0.914	0.1418	−6.450	$< .001$
15		17-1	−1.167	0.151	−7.755	$< .001$
16		17%-1	−1.268	0.157	−8.103	$< .001$

Table 5.15.: Lmer models for looks to the low-frequency object during NP2, Trial Category 1, Exp. 5
$inspections - 50\% \sim VerbType + (1|sub) + (1|item), family = binomial(link = "logit"$

	Predictor	Coef.	SE	Wald z	p
1	(Int.) (Cond. N)	0.034	0.120	0.281	$= .779$
2	Cond. R	−1.080	0.138	−7.822	$< .001$

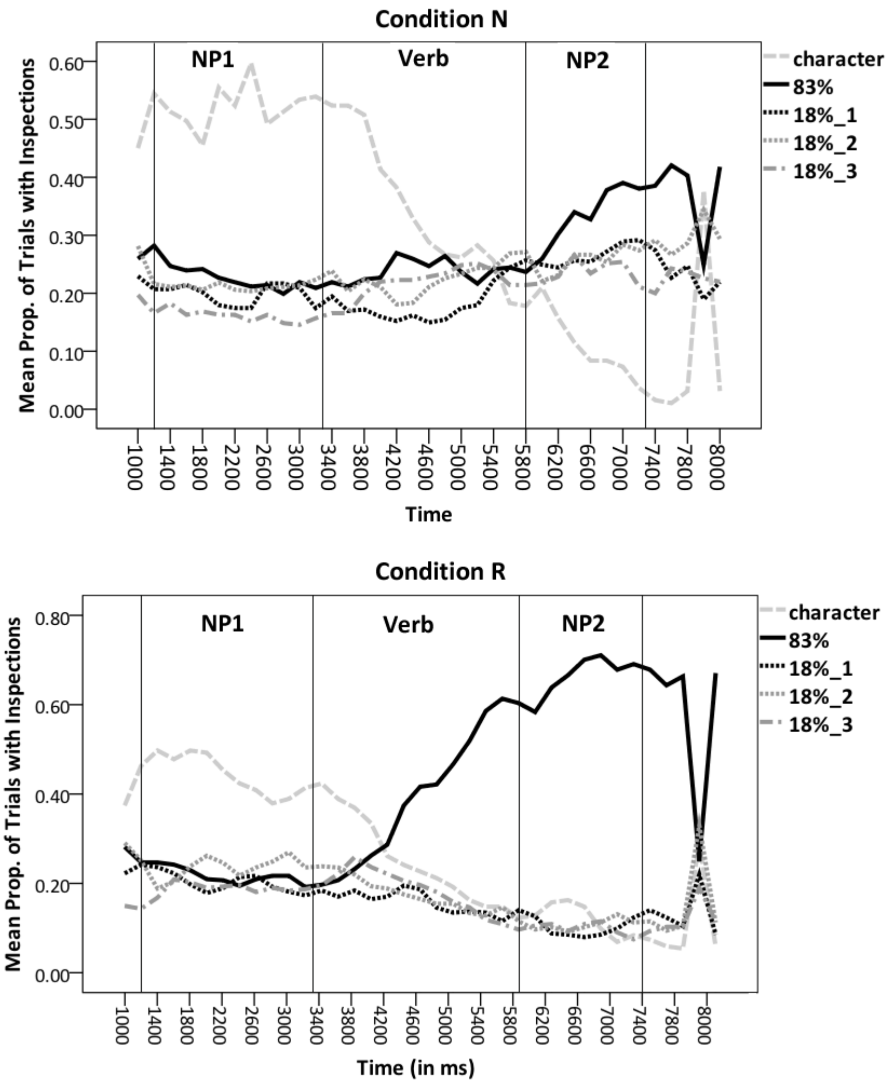


Figure 5.7.: Timegraphs Exp. 5, Trial Category 2, Condition N (top) and Condition R (bottom)

Eye-movements during Verb and NP2 in Trials of Trial Category 2 Eye movements during trials of Category 2 (low-frequency object included) were also analyzed. We found main effects of ROI for verb region in both conditions (N: $\chi(3) = 12.089, p < .01$, Table 5.16, Rows 2-4; R: $\chi(3) = 115.86, p < .001$, Table 5.16, Rows 10-12) and for NP2 in Condition R ($\chi(3) = 64.588, p < .001$, Table 5.16, Rows 14-16). There were no reliable differences in proportions of inspections to different ROIs during NP2 in trials of Condition N, however ($\chi(3) = 5.238, p = .155$, Table 5.16, Rows 6-8). The effects for verb region and NP2 in

Condition R are caused by the high-frequency target being looked at significantly more often than the other objects (i.e., the three distractors). In verb region of N, the high-frequency target was inspected less often than the other objects. See Figure 5.7 for timegraphs and Figure 5.8 for barcharts.

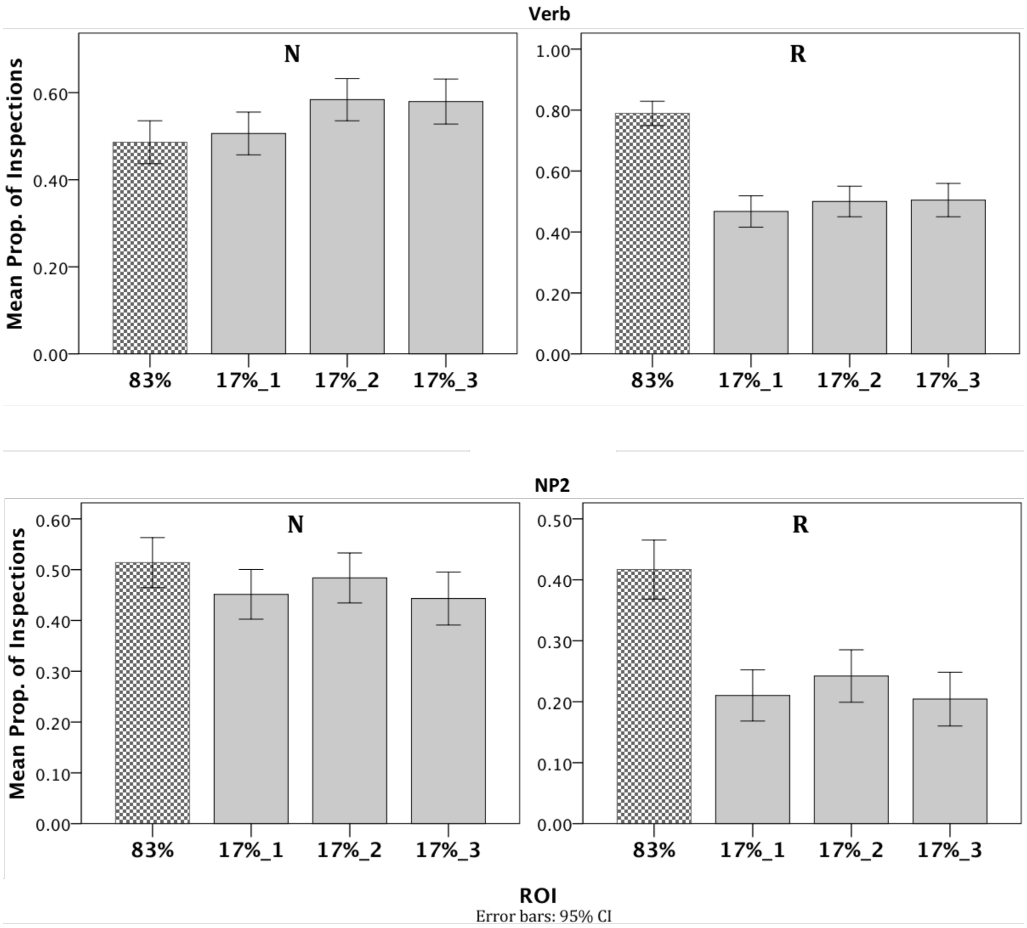


Figure 5.8.: Proportions of trials with at least one inspections to ROIs during the verb (top charts) and NP2 (bottom charts) in trials of Category 2, Day 1, Exp. 5: Condition N (left charts) and Condition R (right charts)

Eye-movements during Verb and NP2 in Trials in Block 2 of Trial Categories 1 and 2

We also ran analyses for eyes movements individually for Block 2, in order to see whether the picture would be clearer once learners got used to the learning scenario. In fact we found some differences for Trial-Type Category 2, Condition N: There was a reliable effect in NP2 now ($\chi(3) = 11.621, p < .01$) and the high-frequency object was inspected reliably more than the other objects (except one distractor). See Table 5.17.

Table 5.16.: Lmer models for inspections on ROIs during the verb and NP2, Trial Type 2, Exp. 5
$Inspections \sim ROI + (1|sub) + (1|item) + (1|block), family = binomial(link = "logit")$

		Predictor	Coef.	SE	Wald z	p
Condition N						
1	*V*	(Int) (high-f.)	−0.063	0.137	−0.460	= .646
2		low-f.	0.068	0.144	0.470	= .639
3		17%-1	0.402	0.145	2.761	< .010
4		17%-1	0.386	0.150	2.572	< .050
5	*NP2*	(Int) (high-f.)	0.067	0.136	0.499	= .618
6		low-f.	−0.268	0.144	−1.869	= .062
7		17%-1	−0.134	0.143	−0.935	= .350
8		17%-1	−0.299	0.149	−2.015	< .050
Condition R						
9	*V*	(Int) (high-f.)	1.3291	0.130	10.243	< .001
10		low-f.	−1.468	0.162	−9.091	< .001
11		17%-1	−1.334	0.160	−8.353	< .001
12		17%-1	−1.320	0.166	−7.958	< .001
13	*NP2*	(Int) (high-f.)	−0.387	0.176	−2.200	< .050
14		low-f.	−1.122	0.170	−6.608	< .001
15		17%-1	−0.926	0.162	−5.705	< .001
16		17%-1	−1.152	0.179	−6.445	< .001

Table 5.17.: Lmer models for inspections on ROIs during the verb and NP2, Trial Category 2, Condition N, Block 2, Exp. 5
$Inspections \sim ROI + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	Wald z	p
1	(Int) (high-f.)	0.118	0.143	0.827	= .408
2	17%-1	−0.594	0.195	−3.046	< .010
3	17%-2	−0.164	0.194	−0.847	= .397
4	17%-3	−0.474	0.203	−2.331	< .050

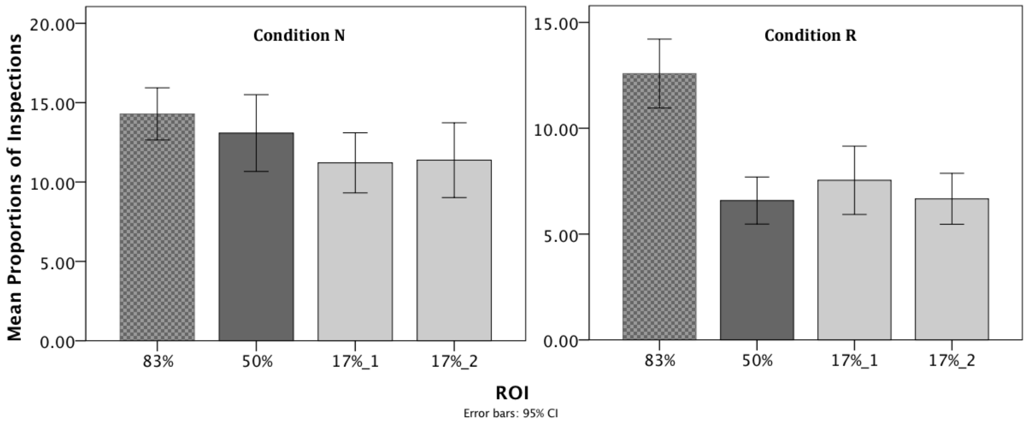


Figure 5.9.: Proportions of trials with at least one inspection to ROIs during trials of Test Type 1 in Conditions N (left chart) and R (right chart), Exp. 5

Eye-movements during Vocabulary Test in Trials of Test Type 1 Finally, we investigated gaze during test trials. On Day 1, for trials of Test Type 1 (83% referent and 50% referent depicted), there were significantly more looks to the high-frequency object than to the three other objects ($\chi(3) = 7.848, p < .05$) and there was an effect of Verb Type (more inspections in Condition N, $\chi(1) = 17.929, p < .001$). We also found a marginal interaction between factors Verb Type and ROI when considering only the high-frequency object and the low frequency object as ROIs ($\chi(1) = 3.052, p = .081$): The difference between both regions was larger in Condition R than Condition N. This marginal interaction was supported by the finding that the difference between the number of inspections to the high-frequency object and the low-frequency object was not significant in Condition N ($\chi(3) = 4.980, p = .173$, Table 5.18, Row 2) but in Condition R ($\chi(3) = 9.540, p < .05$, Table 5.18, Row 2; see Figure 5.9). On Day 2, the pattern of eye-movements was similar except that there was no interaction between ROI and Condition ($\chi(2) = 3.430, p = .180$) and no main effect of ROI in Condition R ($\chi(3) = 4.849, p = .183$).

Eye-movements during Vocabulary Test in Trials of Test Type 2 For trials of Test Type 2 (50% object and category associate (CA) depicted), we found no effect of ROI ($\chi(3) = 4.135, p = .247$) or Condition ($\chi(1) = 1.754, p = .185$) in eye movements, nor an interaction ($\chi(7) = 6.886, p = .441$). Considering only the low-frequency object and the category associate did not change this picture: There was no effect of ROI for Condition N ($\chi(3) = 1.608, p = .658$) nor for R (Figure 5.10; $\chi(3) = 3.359, p = .340$). Eye-movements on Day 2 reflected the same pattern.

Table 5.18.: Lmer models for inspections on ROIs during vocabulary testing, Test Type 1, Day 1, Exp. 5

$$Inspections \sim ROI + (1|sub) + (1|item), family = binomial(link = "logit")$$

	Predictor	Coef.	SE	Wald z	p
Condition N					
1	(Int) (high-f.)	1.636	0.151	10.838	$< .001$
2	low-freq.	−0.066	0.193	−0.340	$= .734$
3	17%-1	−0.361	0.189	−1.917	$= .056$
4	17%-2	−0.292	0.191	−1.534	$= .125$
Condition R					
1	(Int) (high-f.)	1.347	0.134	9.704	$< .001$
2	low-freq.	−0.580	0.191	−3.034	$< .010$
3	17%-1	−0.343	0.193	−1.772	$= .076$
4	17%-2	−0.319	0.201	−1.588	$= .112$

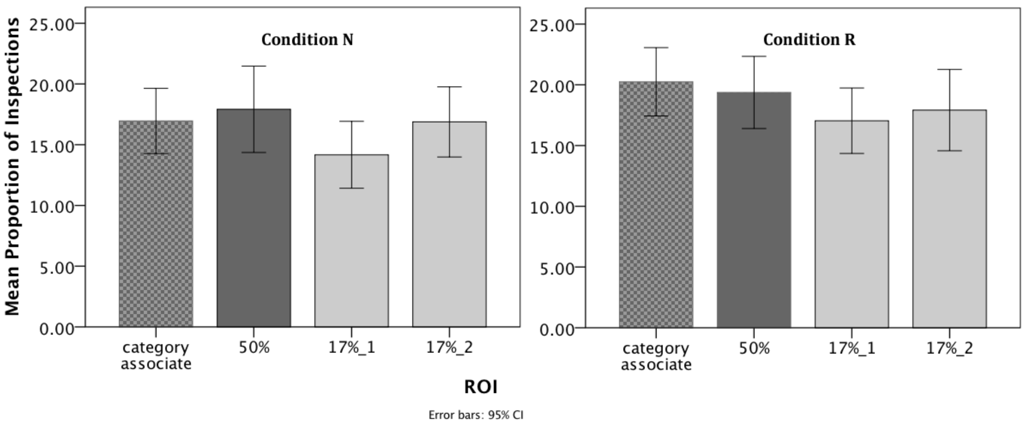


Figure 5.10.: Proportions of trials with at least one inspection to ROIs during trials of Test Type 2 in Conditions N (left chart) and R (right chart), collapsed across days, Exp. 5

Table 5.19.: Lmer models & p-values from MCMC sampling for inspections to low-freq. object during vocabulary testing collapsed across days, Test Type 2, Exp. 5
$Inspections \sim ROI + (1|sub) + (1|item), family = binomial(link = "logit")$

	Predictor	Coef.	SE	t	MCMCmean	pMCMC	$Pr(>\|t\|)$
1	(Int) (Non-R.)	0.286	0.009	31.543	0.281	0.110	$> .001$
2	Restric.	-0.033	0.010	-3.333	-0.033	$> .010$	$> .001$

Eye-movements during Vocabulary Test on Day 1 and Day 2 Since eye-movement data for single trial types and conditions was sparse (8 trials each), we collapsed data from both days. Crucially, we found a clear interaction between factors Verb Type and ROI (high-frequency vs. low frequency competitor) for Test Type 1 ($\chi(1) = 4.035, p < .05$): The difference between looks to high-frequency and low-frequency objects was larger in Condition R than Condition N. For Test Type 2, Verb Type predicted the proportion of looks to the low-frequency competitor: It was reliably higher in Condition N than in Condition R ($\chi(1) = 11.035, p < .001$; Table 5.19).

5.3.5. Summary and Discussion

5.3.5.1. Summary

Learning rates (83% choices in Test Type 1) in both conditions and on both days revealed successful learning. As expected, learning rates were higher in Condition R than in Condition N. Performance was also better on Day 1 than Day 2 but the difference was only significant in Condition N.

Crucially, we found differences between conditions concerning the meanings that participants chose in the vocabulary tests on both days. While in Test Type 1 (85% object and 50% object depicted) the 83% referent was chosen most frequently in both conditions, this preference was significantly stronger in Condition R than N (both days). The 50% referent, on the other hand, was chosen significantly more often in Condition N than in Condition R (both days). On Day 2, the 50% object was also selected significantly more often than the distractors in Condition N. In R-trials of Test Type 2 (50% object and CA depicted), learning rates on both days reveal a clear preference for selecting the category associate (CA).

For N-trials of Test Type 2, on the contrary, we found that while the CA was selected significantly more often than the distractor objects only on Day 1, learners selected the 50% referent significantly more often than the two distractors on both days. The difference between CA choices and 50% choices was not significant on either day. On both days, learners chose the CA significantly more frequently in Condition R than Condition N whereas they picked the 50% referent significantly more often in Condition N than Condition R. Data further reflects a difference between verb types concerning the effect of factor Day on the chosen meanings: Only for Condition N we found both a main effect of Day on Chosen Meaning as well as a decrease in CA choices from Day 1 to Day 2.

Confidence ratings were higher for Test Type 1 than Test Type 2 and for Day 1 than Day 2. In Trial Type 1, learners rated 83% choices higher than 50% choices on both days; R-choices were rated higher than N-choices. In Trial Type 2, in contrast, ratings were higher for Condition N than Condition R and for 50% choices than CA choices. In addition, we found an interaction between factors Condition and Chosen Meaning for Trial Type 2 on Day 1: The difference between confidence ratings for CA choices and 50% choices was reliably larger in Condition R than in Condition N.

Eye-movement data recorded during learning trials of Category 1 reveal that the 83% referent was inspected significantly more often then the other objects during NP2 in both conditions and during the verb in Condition R. The interactions for both time regions between factors Condition and ROI (with the two ROIs 83% object and 50%) show that the difference between looks to the 83% referent and to the 50% referent was more pronounced in Condition R than in Condition N. In R-trials of Category 2, the 83% candidate was looked at significantly more frequently than the three distractors during both the verb and NP2. Unexpectedly, the 83% object was inspected significantly less often than the three distractors during the verb of N-trials of Category 2.

For eye-movements during Test-Type-1 trials, we found a significant interaction between factors ROI (83% object and 50% object) and Condition: While in Condition N, learners looked significantly more often at the 83% referent than at the 50% object, this difference was not reliable in Condition R. The interaction was marginally significant on Day 1 and clearly significant when data of both days was collapsed. The collapsed data moreover revealed a significant difference between conditions regarding the proportion of looks to the 50% object in trials of Test Type 2: It was higher in Condition N than in Condition R.

5.3.5.2. Discussion

The Interaction of CSWL and SLCL Learning rates (83% choices in Test Type 1) in Experiment 5 reveal that participants can be successful in learning nouns based on our paradigm even with the enlarged vocabulary size. The very clear difference in learning performance and confidence ratings in Test Type 1 between Conditions R and N support, again, the view that verbal restrictions (SLCL) result into better learning than CSWL. While in principle, learning in Condition R could (partly) be due to CSWL learning, this is very unlikely given the eye-movements during the verb and NP2 of learning trials (see timegraphs in Figure 5.5): Learners only considered the high-frequency target, ignoring all other objects.

As expected, we found for Test Type 1 (85% object and 50% object depicted) that the high-frequency candidate was chosen most frequently in both conditions and on both days: It was the best candidate according to both CSWL and SLCL. Crucially, we could show that this dominance was clearer in Condition R than in Condition N and, in addition, that confidence ratings were reliably higher in Condition R than Condition N, as expected. We attribute this pattern to the nature of SLCL, which is more direct and therefore more helpful than CSWL: In a setting as in Experiment 5, the target referent can be *immediately* and *unambiguously* identified via SLCL because the learner has a perfect knowledge about edible objects and there is exactly one edible object in the scene. The parallel nature of CSWL was clearly seen on the 50% object, which was chosen reliably more often in Condition N than in Condition R. On Day 2 in Condition N, it was even chosen reliably more often than the distractors, despite the availability of the 'best candidate', that is, the 83% object. This pattern clearly reveals that CSWL-learners were sensitive to the fine-grained difference between the co-occurrence frequency of 17% and the co-occurrence frequency of 50%.

Results from Test Type 2 (50% object and CA depicted) essentially enrich this interpretation. In Condition R, learners clearly preferred to select the category associate and showed no sensitivity for the low-frequency object at all on either day. In Condition N, in contrast, the category associate was preferred over the two distractors only on Day 1 whereas learners chose the 50% object significantly more often than the distractors on both days. This pattern unambiguously reveals a sensitivity for smaller statistical differences between the co-occurrences of objects and nouns (in particular 83% versus 50% and 50% versus 17%) in Condition N, which is completely blocked in Condition R. We attribute the blocking effect to the impact of SLCL: SLCL is deterministic and direct in that it immediately and unambiguously helps learners to identify referents; in addition, SLCL is trustworthy in that it is based on learners' experience about plausibility. Therefore, learners may have completely relied on its information while ignoring fine-grained differences in co-occurrence frequencies between words and referents.

The facts that, in Test Type 2, the category associate was the only preferred object in Condition R and that the category associate was chosen reliably more often in Condition R than Condition N reveal that SLCL increased participants' sensitivity for category memberships.

While we are surprised that the category associate was chosen significantly more often than the distractors on Day 1 in Condition N, the lack of this effect on Day 2 relativizes the finding again. Moreover, as confidence ratings of Day 1 reveal, in Condition N, learners were less confident when they chose the category associate than when they selected the low-frequency candidate, while in Condition R it was the other way round. We suggest that CA-choices that learners made on Day 1 were subject to a spontaneous but not long-lasting strategy of selecting an object which is semantically closest to the non-present 83% referent. This strategy may have even sometimes blocked learners from selecting the 50% referent although they were sensible to its higher co-occurrence frequencies compared to the 17%-distractors. This additionally indicates how different learning mechanisms and strategies can interact.

Eye-movements support both the sensitivity for fine-grained differences in co-occurrences, which was present in Condition N but blocked in Condition R, and the enhanced sensitivity for category association in Condition R: While there was a very clear preference during learning trials of Category 1 (83% object and 50% object depicted) to inspect the high-frequency candidate in Condition R, both the high-frequency object and low-frequency object were considered in Condition N. The preference for inspecting the high-frequency object in learning trials of Category 2 was also more pronounced in Condition R than Condition N. For gaze during vocabulary-test trials of Type 1, we found an interaction between factors Verb Type and Chosen Object (83% vs. 50%), revealing a clear preference for the high-frequency object in Condition R and a secondary preference for the low-frequency object in Condition N. Gaze during trials of Test Type 2 did not provide a clear picture. Seemingly, the higher overall uncertainty in these trials (as reflected by lower confidence ratings in Test Type 2 than Test Type 1) caused learners to compare a lot between objects and hence to make more inspections everywhere. Perhaps unsurprisingly, the amount of inspections to any ROI was significantly higher in Test Type 2 than in Test Type 1($\chi(1) = 29.051, p < .001$).

As predicted, we also found anticipatory eye-movements to objects which match verbal restrictions in verb region of learning trials in Condition R. Surprisingly, however, we additionally found that the high-frequency object was looked at significantly less than the distractors in verb region of Condition N. This is rather puzzling because the verb provided no information, the order of trials was random, and there was no visual dominance of any depicted object in any way. We therefore ascribe this effect to noise. Since Experiment 5 was an even more challenging task for learners than Experiments 1 to 4, with even more nouns to be learned, it is perfectly possible that other mental processes than the ones under investigation influenced learners' eye movements. We are nonetheless confident with our interpretations, as they are supported by the clear patterns in our off-line results.

Consolidation It is interesting and encouraging that noun learning was still above chance on Day 2 for both conditions, demonstrating that learning resulting from our paradigm can be long-lasting. Furthermore, there were differences in the vocabulary-test data between both days: As expected, learning rates and confidence ratings decreased from Day 1 to Day 2. The result that, in Test Type 1 in Condition N, the low-frequency candidate was chosen more often on Day 2 than on Day 1, may reveal that there was a small loss in sensitivity for co-occurrence differences (83% and the 50%) on Day 2. This may have been caused by a memory decay concerning co-occurrence relations. On the other hand, the amount of 50% choices in Condition N on Day 2 did not decrease. We cannot draw final conclusions about this issue.

The findings that the amount of category-associate choices (Test Type 2) was stable over Days in Condition R and that it decreased in Condition N seems to only partly confirm our hypothesis regarding over-night generalization of semantic categories: Semantic categories did play a greater role when nouns were learned based on verbal constraints but we cannot see an increase of generalization over days. This missing rise may be due to a ceiling effect in Condition R or it may falsify our hypothesis that consolidation enhances generalization on a conceptual level. An alternative interpretation of the data is that consolidation in fact caused a *decrease* rather than an increase in sensitivity to category membership which, however, was blocked by the enhancement of sensitivity to category membership in Condition R. That is, if there was no focus on category associations during noun learning, as in pure CSWL learning (Condition N), learners did not store noun meanings based on these associations and could therefore not recall categories well in the test on Day 2.

To summarize our main results, we found clear evidence that CSWL works in a parallel and probabilistic manner (i.e., different potential referents and their probabilities concerning co-occurrences with a noun are tracked in parallel) while SLCL is more deterministic and direct in nature (i.e., only one candidate is, or few candidates are, immediately identified and considered as referents). Our data further reveals that SLCL can completely block CSWL representations of low-frequency meanings when both mechanisms are independently applicable. We also found that SLCL increases learners' sensitivity for category associations. Our data also reflects that learning which results from our paradigm is potentially long-lasting. Finally, results from Experiment 5 suggest that over-night consolidation may cause a small loss of sensitivity for differences in co-occurrence frequencies and a decrease in sensitivity to category associations, which, however, may be blocked when the applied learning mechanism increases this sensitivity.

6. General Discussion

In this thesis, the interplay of two visually situated mechanisms of second-language word learning in non-instructed and multi-modal environments was investigated, cross-situational word learning (CSWL) and learning based on sentence-level constraints (SLCL). As laid out in Chapter 2, although both mechanisms have been shown to help adult word learners in both more implicit and more explicit learning situations, their interplay has so far not been addressed. We considered it an important enterprise to contribute to closing this gap for two major reasons: Out of the diverse range of word-learning mechanisms, CSWL and SLCL are most clearly relevant for adults in non-instructed environments. Moreover, we expected the interaction of CSWL and SLCL to be potentially insightful due to their differences in nature: While CSWL is a bottom-up associative process, SLCL offers top-down constraints on meaning. We further hypothesized CSWL to be parallel, incremental, and probabilistic and SLCL to be more deterministic. In order to provide evidence for our working hypothesis that word learning in adults in non-instructed environments is interactive, multi-modal, and constraint-based, we conducted a series of five step-wise learning experiments: In a first phase, participants learned a set of verbs, in a second step they were trained on nouns, and in a last step, noun knowledge was assessed. Crucially, our experimental paradigm was characterized by integrating a naturalized learning setting, compared to previous word acquisition studies: Nouns were embedded as parts of sentences and sentences were situated within visual scenes.

In Experiment 1, we evaluated the use of CSWL and SLCL (i.e., verbal constraints together with the visual context and learners' world knowledge) for noun learning in this naturalized situation. This was done in order to find out, firstly, whether CSWL operates successfully, secondly, whether SLCL boosts noun learning and, thirdly, whether language novices process sentences multi-modally and incrementally and, specifically, exploit verbal restrictions to predictively identify referents on-line. To investigate the potential boost of SLCL, we compared subject-noun learning performance to object-noun learning performance. Results reveal that learners applied both CSWL and SLCL in a complementary way to learn the vocabulary and that they processed the sentences multi-modally, incrementally, and predictively. However, while SLCL (verbal constraints) clearly boosted on-line identification of referents, the enhancement of noun learning was not entirely clear (object-noun learning was not reliably better than subject-noun learning). Finally, data suggests a potential positive effect of predicting referents for post-verbal nouns on noun learning (as evidenced by a tendency for the proportion of looks to the target object during the verb to correlate with object-name learning).

In Experiment 2, we took this last issue up and directly investigated the effect of verb-based prediction on noun learning. In order to follow this aim, we introduced a second word order (OVS), which is characterized by a verb which follows rather than precedes the object noun. Besides expecting to replicate findings from Experiment 1 and find CSWL for a second word order, we predicted that object nouns would be learned better in SVO than OVS whereas subject nouns would be learned equally well in both conditions. While we found that learners in the SVO condition outperformed learners in the OVS condition, the expected interaction between condition and noun type was not reliable. Eye-movements gave further cause for careful interpretation: OVS-learners may have suffered from word-order based uncertainty. Results are, however, at least in accordance with the hypothesis that verb-based prediction of referents has a positive influence on noun learning.

In Experiment 3, we re-addressed the question whether SLCL boosts noun learning. We also examined the interaction of CSWL and SLCL when both are applicable in a complementary way with regard to identifying the meaning of one word. To achieve this objective, we manipulated the degree of referential uncertainty within three conditions: Either nouns followed restrictive verbs and were accompanied by scenes which depicted only one object matching verbal constraints, or verbs were restrictive but scenes contained two objects belonging to the category required by the verb, or nouns followed non-restrictive verbs which means that all four depicted objects matched the verb semantically. On-line and off-line results provide evidence for the hypotheses that, firstly, SLCL boosts noun learning and secondly, SLCL and CSWL interact in that they can be jointly used to identify a noun's meaning. More specifically, we found that CSWL can compensate for a low degree of referential uncertainty that remains based on SLCL.

Experiment 4 was conducted in order to continue investigating the interaction of CSWL and SLCL, specifically, when both mechanisms are in conflict. We additionally evaluated the use of CSWL and SLCL in a further naturalized learning setting. In particular, the design of Experiment 4 was characterized by nouns which had two potential meanings (i.e., two potential referents) - the high-frequency referent co-occurred with the noun in 83% of the noun's presentations, the low-frequency candidate had a co-occurrence rate of 50%. Each noun was in one of two conditions: Either it always followed a non-restrictive verb (Condition N) or it was preceded by a restrictive verb half of the time (Condition R). CSWL supported the 83% meaning in both conditions. Crucially, the restrictive verb in Condition R, on the contrary, supported the 50% meaning which means that SLCL and CSWL conflicted. Because those trials which contained a restrictive verb (e.g., *eat*) in Condition R were exactly those trials which did not contain the SLCL-supported 50% referent (i.e., something edible), referential uncertainty was two-sided in Experiment 4. While learning rates clearly reveal that CSWL and SLCL were evenly influential with regard to learners' decisions in the vocabulary test (50% referents and 83% referents were chosen equally often in Condition R), eye-movements suggest that CSWL was slightly dominant by still influencing learners' attention on-line even if they chose the SLCL-supported meaning.

The aim of Experiment 5 was two-fold: Firstly, in order to examine the nature of both mechanisms, we studied the interaction of CSWL and SLCL in a scenario where CSWL and SLCL were independently applicable. Secondly, we included an assessment of vocabulary one day after learning to evaluate whether CSWL and SLCL also give rise to long-lasting word learning. As is Experiment 4, nouns had two potential meanings (again, there was an 83% referent and a 50% referent) and they were in one of two conditions (R or N). This time, however, restrictive verbs and matching referent were co-present in Condition R. Crucially, both CSWL and SLCL supported the 83% candidate. There were two different types of trials in the four-forced-choice vocabulary test, either both the 83% object and the 50% object as well as two distractors were selectable (Type 1) or only the 50% candidate was available, besides a category associate (an object of the same semantic category as the missing 83% referent) and two distractors (Type 2). We found that for Test Type 1, learners were more likely to select the 83% referent in Condition R than in Condition N and more likely to select the 50% candidate in Condition N than in Condition R. Moreover, in Test Type 2, the 50% referent was preferred over the distractors (which had a co-occurrence rate of 17%) in Condition N but not in Condition R. In Condition R, on the contrary, there was a clear preference to choose the category associate. This pattern of results clearly provides evidence for the hypotheses that CSWL learning elicited a sensitivity to smaller co-occurrence frequencies (83% versus 50% and 50% versus 17%) which on the contrary was completely blocked when SLCL was available. SLCL, in contrast, increased learners' sensitivity to category membership. Eye-movements during training and testing support these interpretations. Results from the vocabulary test on Day 2 additionally reveal that learning rates in both conditions were still clearly above chance and only significantly worse than on Day 1 in Condition N. We found that there was a small loss in sensitivity for co-occurrence differences (83% and the 50%) in Condition N on Day 2 which we attributed to a potential memory decay concerning co-occurrence relations.

6.1. The Interaction of Word-Learning Mechanisms

Our experiments shed light on the interplay of CSWL and SLCL: Both mechanisms can be used in a complementary way, either to learn a set of nouns, or to learn a single noun meaning; CSWL and SLCL influence word learning about equally strongly when they are in conflict; SLCL blocks CSWL when both mechanisms are independently (i.e., redundantly) applicable. We additionally found that SLCL is used whenever it is available (for object nouns in Experiment 1 and in Condition SVO of Experiment 2, in Conditions NoRu and LowRU of Experiment 3, and in Condition R of Experiments 4 and 5). This, on the contrary, is not the case for CSWL, which is only applied when SLCL is either not available at all (subject nouns in Experiments 1 and 2, Condition HighRU of Experiment 3, Condition N of Experiments 4 and 5) or not informative enough (Condition LowRu of Experiment 3), and when SLCL and CSWL are in conflict (Condition R of Experiment 4).

This pattern suggests that applying CSWL and SLCL in parallel is gainful for learners when both mechanisms complement each other. When learners have the choice because information is redundant, they rely more strongly on SLCL than CSWL. The reason for this behavior most likely is that SLCL is a more helpful cue than CSWL in that it can unambiguously identify referents. Additionally, SLCL operates based on learners' solid word knowledge. Interestingly, the finding that CSWL is ignored when it is redundant implies that conducting CSWL does not happen automatically and effortlessly but that it causes processing cost which the learner saves when possible. However, when SLCL and CSWL are in conflict, learners use both mechanisms rather than ignoring CSWL, which indicates that SLCL is not given general priority. Rather, learners must in some way differentiate between situations in which they can ignore co-occurrence frequencies (CSWL) and situations in which they cannot - because they need the information or because there is a conflict. We therefore propose a model of an efficient, 'lazy but careful' adult word learner who employs as many resources in parallel as necessary but ignores the less direct and helpful cue when information is redundant. However, when the relevance of different cues is unclear, he considers all of them. This bounded-parallel like strategy of word learning may of course also be problematic in case the relevance of cues has been mis-judged, that is, when information which has been discarded becomes relevant later again. While this account is generally in accordance with existing computational models of word learning which integrate more than one learning cue (Yu & Ballard, 2007; Alishahi & Fazly, 2010; Frank, Goodman & Tenenbaum, 2009), it is much more detailed regarding the precise ways word-learning mechanisms interact.

6.2. Applications

While our results most directly allow us to draw conclusions for a theory of word learning in adults, there are additionally some implications for applications that can be drawn. As mentioned in Chapter 2, statistical word learning has been shown to be a very useful tool for adult vocabulary learning and for re-learning of words in aphasics (Breitenstein and colleagues). One characteristic that potentially makes statistical word learning a good way of learning is the fact that it is based on repeated exposure (see also Blythe et al., 2010).

> 'Our findings demonstrate that high intensity active exposure is a powerful mechanism of adult language acquisition.' (Breitenstein et al., 2004: 455)

It is also extremely easy to implement because minimally it only requires presenting spoken or written words and pictures concurrently, for instance on a computer screen or on cards. A crucial question, however, is which degree of task complexity is ideal, that is, which degree of task complexity makes learning sufficiently interesting and appealing while not being too demanding. Breitenstein and colleagues (Breitenstein et al. 2004; 2007; Dobel et al. 2010; Breitenstein, p.c.) achieved good performance with healthy and aphasic adults, using a very easy interactive method: At each trial, participants were asked to decide whether one picture and one word matched or not by pressing a button. Performance after five days improved enormously in all experiments (up to 94% of 45 words). While this is an excellent result, results from our experiments, using more demanding tasks with interacting word-learning mechanisms, are potentially even more promising. On Day 2 of Experiment 5, that is, after only 40 minutes of training one day before, learners still retained 78% of the eight nouns in Condition R (and 63% of all 16 nouns). It would be worthwhile to conduct an experiment over more days and including repeated training to see how many words can potentially be learned how quickly using our paradigm. In fact, most of our about 200 participants were very engaged and motivated and often reported after the experiments that they had enjoyed it. This suggests that some degree of challenge might be desirable and that our paradigm has potential as a vocabulary learning tool.

While learning as many words as possible in as little time as possible is one possible aim of a vocabulary tool, it is not the only one. Another goal can be, for instance, to learn words as deeply as possible or in all ways they can be used. Given that this last objective naturally requires presenting words in different linguistic contexts, our paradigm would be more appropriate than learning environments as in Breitenstein et al.'s experiments.

There are other existing vocabulary learning tools which employ both linguistic contexts (and sometimes discourses) and visual information, pictures or videos, but they use off-line and extremely explicit glosses rather than more natural accompanying scenes (e.g., Plass, Chun, Mayer & Leutner, 1998; Jones & Plass, 2002; Al-Seghayer, 2001; Yanguas, 2009; Yeh & Wang, 2003; Jones, 2006). The vocabulary trainer by Jones (2006), for instance (Figure 6.1), is based on graphical user interfaces which depict written words or phrases. When learners do not understand these words, they can additionally request pictures, pronunciation records, or linguistic contexts by clicking on corresponding symbols. The advantage of our more natural and on-line presentation of combined multi-modal information may be, firstly, that the task is more interesting because it is more game-like.

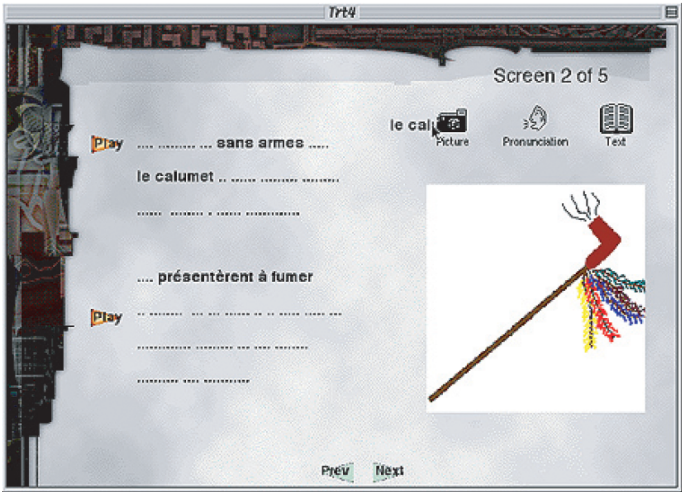


Figure 6.1.: Vocabulary trainer by Jones (2006): Example trial

Secondly, it may teach language novices skills that they can apply in real-life situations, specifically, to better exploit information from both the spoken linguistic context and the visual environment.

Finally, our results suggest a number of smaller aspects which are potentially relevant for the development of vocabulary trainers. Firstly, we experienced that it is relevant whether words (or sentences) are presented in blocks or not: For noun learning, for instance, we found that learners were more successful when only two to four nouns were repeated within one block than when all words were randomly presented. Tightly combined with this factor is the question about the best time lag between repetitions. We experienced that CSWL was difficult for participants when there were more than four to five trials between exposures of a noun. While the absolute number of exposures was another relevant factors, it was not the most important.

6.3. Future Research

Given that there are several further learning mechanisms that we have not examined in this series of experiments, our argument that word learning is interactive enforces the question how the interplay of CSWL and SLCL with these other mechanisms may be like. One possible direction for future experiments is to integrate social cues into our learning paradigm. We expect certain social cues, namely gaze and pointing gesture, to behave in a similarly deterministic way as SLCL. Eye gaze of an interlocutor is a potentially valuable cue because speakers tend to look at something shortly before they start to speak about it (Meyer, Sleiderink& Levelt, 1998). Children have been shown to exploit this regularity to find the meanings of novel words (Baldwin, 1991). Pointing gestures can also uniquely identify referents of novel words. Since people learn how to interpret these social cues from childhood on it is likely that they are even stronger and more direct than sentence-level constraints. It would be interesting, for instance, how SLCL interacts with gaze, whether gaze overrides SLCL in conflicting situations and whether both cues are redundant.

Another interesting objective would be to expand the learning procedure over more days in order to investigate the scope of the paradigm, for instance, compared to Breitenstein et al.'s application. For direct evaluation of the paradigm as a learning tool, it would moreover be necessary to conduct more comparisons concerning the complexity of the learning task (number of new words per training unit or block, number of exposures, manipulation of order, degree of referential uncertainty, etc.).

6.4. Conclusions

To summarize, our main contributions are the following findings: CSWL and SLCL (based on verbal constraints, the visual context, and people's world knowledge) are powerful mechanisms for adult (L2) word learners in non-instructed and multi-modal environments, potentially leading into long-lasting retention; adult second language learners can process simple spoken sentences in a native-like way, in particular, they can rapidly, incrementally, and even predictively integrate multi-modal information into their comprehension; SLCL immediately and reliably helps language novices to identify visual referents on-line, effectively compensating for comprehension gaps; SLCL and CSWL can jointly be used to learn a set of nouns but also to learn the meaning of a single noun; when CSWL and SLCL are in conflict, both are considered about equally strongly, with a slight dominance of CSWL on an attentive level; while CSWL is a parallel, probabilistic and bottom-up way of learning, SLCL works more deterministic, category-based, and top-down; SLCL can completely block sensitivity to differences in co-occurrence frequencies that otherwise characterize CSWL; word learning in adults is an interactive and constraint-based process as reflected by the multiple interactions resulting from our experiments. These results suggest that while adult word learners apply different learning mechanisms in parallel if necessary, they only consider the most direct cue when information is redundant; however, they do not ignore any conflicting information.

Our contributions are insightful for a theory of second-language word learning and they support the validity of an interactive constraint-based word learning (and processing) model. Furthermore, they strongly suggest that it is important to study word-learning mechanisms in a situated and integrating way. Finally, the research presented in this thesis demonstrates that it can be gainful for language learning researchers to use combined and on-line methods.

A. Experimental Items Experiments 1 & 2

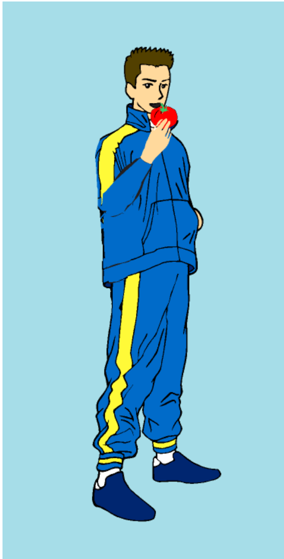


Figure A.1.: Example picture verb learning/testing, Exp. 1/Exp. 2
bermamema
'eat'

Table A.1.: Verb-Learning Items Exp. 1

Item	Verb	Translation	Depicted Character	Depicted Object
1	bermamema	eat	woman	chicken
2	mankemema	barbecue	woman	tomato
3	bumbumema	salt	woman	tomato
4	rupamema	dry	woman	shorts
5	melimema	iron	woman	scarf
6	tambamema	sew	woman	shorts
7	bermamema	eat	man	tomato
8	mankemema	barbecue	man	chicken
9	bumbumema	salt	man	chicken
10	rupamema	dry	man	scarf
11	melimema	iron	man	shorts
12	tambamema	sew	man	scarf

Table A.2.: Verb-Testing Items Exp. 1

Item	Verb	Translation	Depicted Character	Depicted Object
1	bermamema	eat	woman	tomato
2	mankemema	barbecue	woman	chicken
3	bumbumema	salt	woman	chicken
4	rupamema	dry	woman	scarf
5	melimema	iron	woman	shorts
6	tambamema	sew	woman	scarf
7	bermamema	eat	man	chicken
8	mankemema	barbecue	man	tomato
9	bumbumema	salt	man	tomato
10	rupamema	dry	man	shorts
11	melimema	iron	man	scarf
12	tambamema	sew	man	shorts

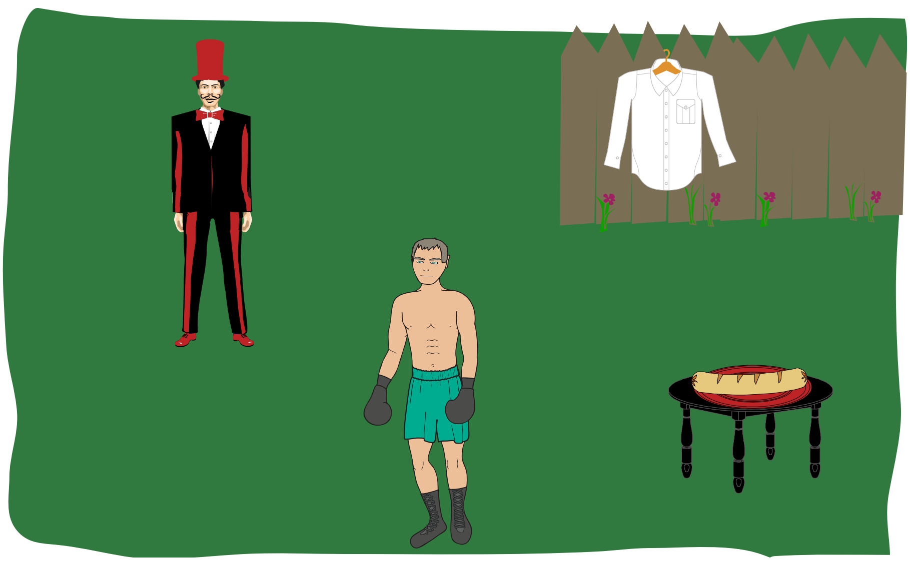


Figure A.2.: Example picture noun learning, Exp. 1/Exp. 2
Si prijas bermamema si sonis
'The director eats the corn.'

Table A.3.: Noun-learning sentences Exp. 1

Item	Sentence	Translation	Scene
1	Si badut bermamema si sonis	The clown eats the sausage	clown, acrobat, sausage, t-shirt
2	Si prijas bermamema si jagung	The director eats the corn	director, magician, corn, shirt
3	Si wanit bermamema si worel	The ballerina eats the mushroom	ballerina, boxer , mushroom, trousers
4	Si prijas bermamema si sonis	The director eats the sausage	director, boxer, sausage , shirt
5	Si wanit bermamema si jagung	The ballerina eats the corn	ballerina , acrobat, corn, trousers
6	Si badut bermama si worel	The clown eats the mushroom	clown, magician, mushroom, t-shirt
7	Si wanit mankemema si sonis	The ballerina barbecues the sausage	ballerina, magician, sausage, trousers
8	Si badut mankemema si jagung	The clown barbecues the corn	clown, boxer, corn, t-shirt
9	Si prijas mankemema si worel	The director barbecues the mushroom	director, acrobat, mushroom, shirt
10	Si badut mankemema si worel	The clown barbecues the mushroom	clown, acrobat, mushroom, shirt
11	Si prijas mankemema si sonis	The director barbecues the sausage	clown, magician, sausage, trousers
12	Si wanit mankemema si jagung	The ballerina barbecues the corn	ballerina, boxer, corn, t-shirt
13	Si prijas bumbumema si jagung	The director salts the corn	director, boxer, corn, trousers
14	Si wanit bumbumema si worel	The ballerina salts the mushroom	ballerina, acrobat, mushroom, t-shirt
15	Si badut bumbumema si sonis	The clown salts the sausage	clown, magician, sausage, shirt
16	Si wanit bumbumema si sonis	The ballerina salts the sausage	ballerina, magician, sausage, t-shirt
17	Si badut bumbumema si jagung	The clown salts the corn	clown, boxer, corn, shirt
18	Si prijas bumbumema si worel	The director salts the mushroom	director, acrobat, mushroom, trousers
19	Si daram rupamema si oblung	The acrobat dries the t-shirt	acrobat, director, t-shirt, sausage
20	Si cowok rupamema si kemei	The magician dries the shirt	magician, clown, shirt, corn
21	Si petin rupamemema si zelan	The boxer dries the trousers	boxer, ballerina, trousers, mushroom
22	Si cowok rupamema si oblung	The magician dries the t-shirt	magician, ballerina, t-shirt, corn
23	Si petin rupamema si kemei	The boxer dries the shirt	boxer, director, shirt, mushroom
24	Si daram rupamema si zelan	The acrobat dries the trousers	acrobat, clown, trousers, sausage
25	Si petin melimema si oblung	The boxer irons the t-shirt	boxer, clown, t-shirt, mushroom
26	Si daram melimema si kemei	The acrobat irons the shirt	acrobat, ballerina, shirt, sausage
27	Si cowok melimema si zelan	The magician irons the trousers	magician, director, trousers, corn
28	Si daram melimema si oblung	The acrobat irons the trousers	acrobat, director, trousers, sausage
29	Si cowok melimema si oblung	The magician irons the t-shirt	magician, clown, t-shirt, corn
30	Si petin melimema si kemei	The boxer irons the shirt	boxer, ballerina, shirt, mushroom
31	Si cowok tambamema si zelan	The magician sews the trousers	magician, ballerina, trousers, corn
32	Si petin tambamema si oblung	The boxer sews the t-shirt	boxer, director, t-shirt, mushroom
33	Si daram tambamema si kemei	The acrobat sews the shirt	acrobat, clown, shirt, sausage
34	Si petin tambamema si zelan	The boxer sews the trousers	boxer, clown, trousers, mushroom
35	Si daram tambamema si oblung	The acrobat sews the t-shirt	acrobat, ballerina, t-shirt, sausage
36	Si cowok tambamema si kemei	The magician sews the shirt	magician, director, shirt, corn

Figure A.3.: Example picture noun testing, Exp. 1/Exp. 2
kowok
'magician'

Table A.4.: Vocabulary-Test Items Exp. 1

Item	Noun	Translation	Depicted Objects
1	sonis	sausage	sausage, corn, mushroom, clown
2	jagung	corn	corn, mushroom, shirt, trousers
3	worel	mushroom	mushroom, sausage, magician, ballerina
4	oblung	t-shirt	t-shirt, trousers, shirt, mushroom
5	zelan	trousers	trousers, shirt, sausage, corn
6	kemei	shirt	shirt, t-shirt, mushroom, acrobat
7	badut	clown	clown, acrobat, corn, sausage
8	wanit	ballerina	ballerina, magician, trousers, t-shirt
9	prijas	director	director, boxer, clown, shirt
10	petin	boxer	boxer, clown, shirt, t-shirt
11	kowok	magician	magician, ballerina, mushroom, trousers
12	daram	acrobat	acrobat, director, corn, sausage

Table A.5.: Sentence Judgement Items Exp. 1

Item	Sentence	Translation	Plausible?	Cause
1	Si badut bermama si jagung.	The clown eats the corn.	yes	
2	Si wanit mankemema si worel.	The ballerina barbecues the mushroom.	yes	
3	Si prijas bumbumema si sonis.	The director salts the sausage.	yes	
4	Si daram rupamema si kemei.	The acrobat dries the shirt.	yes	
5	Si petin melimema si zelan.	The boxer irons the trousers.	yes	
6	Si cowok tambamema si oblung.	The magician sews the t-shirt.	yes	
7	Si petin bermamema si worel.	The magician eats the mushroom.	yes	
8	Si daram mankemema si sonis.	The acrobat barbecues the sausage.	yes	
9	Si cowok bumbumema si jagung.	The magician salts the corn.	yes	
10	Si badut rupamema si kemei.	The clown dries the shirt.	yes	
11	Si wanit melimema si oblung.	The ballerina irons the t-shirt.	yes	
12	Si prijas tambamema si zelan.	The director sews the trousers.	yes	
13	Si badut bumbumema si petin.	The clown salts the boxer.	no	object is character
14	Si wanit bermamema si zelan.	The ballerina eatsthe trousers.	no	clothing instead of food
15	Si prijas mankemema si daram.	The director barbecues the acrobat.	no	object is character
16	Si daram rupamema si sonis.	The acrobat dries the sausage.	no	food instead of clothing
17	Si cowok tambamema si worel.	The magician sews the mushroom.	no	food instead of clothing
18	Si petin melimema si badut.	The boxer irons the clown.	no	object is character
19	Si oblung bumbumema si sonis.	The t-shirt salts the sausage.	no	subject is clothing
20	Si zelan bermamema si worel.	The trousers eats the mushroom.	no	subject is clothing
21	Si kemei mankemema si jagung.	The shirt barbecues the corn.	no	subject is clothing
22	Si sonis melimema si kemei.	The sausage irons the shirt.	no	subject is food
23	Si worel tambamema si zelan.	The mushroom sews the trousers.	no	subject is food
24	Si jagung rupamema si oblung.	The corn dries the t-shirt.	no	subject is food

Table A.6.: Verb-Testing Items Exp. 2

Item	Verb	Translation	Depicted Character	Depicted Object
1	bermama	eat	woman	tomato
2	mankema	barbecue	woman	chicken
3	gumbuma	salt	woman	chicken
4	kipama	dry	woman	scarf
5	felima	iron	woman	shorts
6	tambama	sew	woman	scarf
7	bermama	eat	man	chicken
8	mankema	barbecue	man	tomato
9	gumbuma	salt	man	tomato
10	kipama	dry	man	shorts
11	felima	iron	man	scarf
12	tambama	sew	man	shorts

Table A.7.: Noun-learning sentences Exp. 2, Condition SVO

Item	Sentence	Translation	Scene
1	Si badut bermama si sonis	The clown eats the sausage	clown, acrobat, sausage, t-shirt
2	Si pripas bermama si jafek	The director eats the corn	director, magician, corn, shirt
3	Si gamit bermama si firel	The ballerina eats the mushroom	ballerina, boxer , mushroom, trousers
4	Si pripas bermama si sonis	The director eats the sausage	director, boxer, sausage , shirt
5	Si gamit bermama si jafek	The ballerina eats the corn	ballerina , acrobat, corn, trousers
6	Si badut bermama si firel	The clown eats the mushroom	clown, magician, mushroom, t-shirt
7	Si gamit mankema si sonis	The ballerina barbecues the sausage	ballerina, magician, sausage, trousers
8	Si badut mankema si jafek	The clown barbecues the corn	clown, boxer, corn, t-shirt
9	Si pripas mankema si firel	The director barbecues the mushroom	director, acrobat, mushroom, shirt
10	Si badut mankema si firel	The clown barbecues the mushroom	clown, acrobat, mushroom, shirt
11	Si pripas mankema si sonis	The director barbecues the sausage	clown, magician, sausage, trousers
12	Si gamit mankema si jafek	The ballerina barbecues the corn	ballerina, boxer, corn, t-shirt
13	Si pripas gumbuma si jafek	The director salts the corn	director, boxer, corn, trousers
14	Si gamit gumbuma si firel	The ballerina salts the mushroom	ballerina, acrobat, mushroom, t-shirt
15	Si badut gumbuma si sonis	The clown salts the sausage	clown, magician, sausage, shirt
16	Si gamit gumbuma si sonis	The ballerina salts the sausage	ballerina, magician, sausage, t-shirt
17	Si badut gumbuma si jafek	The clown salts the corn	clown, boxer, corn, shirt
18	Si pripas gumbuma si firel	The director salts the mushroom	director, acrobat, mushroom, trousers
19	Si daram kipoma si oblung	The acrobat dries the t-shirt	acrobat, director, t-shirt, sausage
20	Si towok kipoma si kemei	The magician dries the shirt	magician, clown, shirt, corn
21	Si metin kipomema si zelan	The boxer dries the trousers	boxer, ballerina, trousers, mushroom
22	Si towok kipoma si oblung	The magician dries the t-shirt	magician, ballerina, t-shirt, corn
23	Si metin kipoma si kemei	The boxer dries the shirt	boxer, director, shirt, mushroom
24	Si daram kipoma si zelan	The acrobat dries the trousers	acrobat, clown, trousers, sausage
25	Si metin felima si oblung	The boxer irons the t-shirt	boxer, clown, t-shirt, mushroom
26	Si daram felima si kemei	The acrobat irons the shirt	acrobat, ballerina, shirt, sausage
27	Si towok felima si zelan	The magician irons the trousers	magician, director, trousers, corn
28	Si daram felima si oblung	The acrobat irons the trousers	acrobat, director, trousers, sausage
29	Si towok felima si oblung	The magician irons the t-shirt	magician, clown, t-shirt, corn
30	Si metin felima si kemei	The boxer irons the shirt	boxer, ballerina, shirt, mushroom
31	Si towok tambama si zelan	The magician sews the trousers	magician, ballerina, trousers, corn
32	Si metin tambama si oblung	The boxer sews the t-shirt	boxer, director, t-shirt, mushroom
33	Si daram tambama si kemei	The acrobat sews the shirt	acrobat, clown, shirt, sausage
34	Si metin tambama si zelan	The boxer sews the trousers	boxer, clown, trousers, mushroom
35	Si daram tambama si oblung	The acrobat sews the t-shirt	acrobat, ballerina, t-shirt, sausage
36	Si towok tambama si kemei	The magician sews the shirt	magician, director, shirt, corn

Table A.8.: Noun-learning sentences Exp. 2, Condition OVS

Item	Sentence	Translation	Scene
1	Si sonis bermama si badut	The clown eats the sausage	clown, acrobat, sausage, t-shirt
2	Si jafek bermama si pripas	The director eats the corn	director, magician, corn, shirt
3	Si firel bermama si gamit	The ballerina eats the mushroom	ballerina, boxer , mushroom, trousers
4	Si sonis bermama si pripas	The director eats the sausage	director, boxer, sausage , shirt
5	Si jafek bermama si gamit	The ballerina eats the corn	ballerina , acrobat, corn, trousers
6	Si firel bermama si badut	The clown eats the mushroom	clown, magician, mushroom, t-shirt
7	Si gamit mankema si sonis	The ballerina barbecues the sausage	ballerina, magician, sausage, trousers
8	Si badut mankema si jafek	The clown barbecues the corn	clown, boxer, corn, t-shirt
9	Si pripas mankema si firel	The director barbecues the mushroom	director, acrobat, mushroom, shirt
10	Si badut mankema si firel	The clown barbecues the mushroom	clown, acrobat, mushroom, shirt
11	Si pripas mankema si sonis	The director barbecues the sausage	clown, magician, sausage, trousers
12	Si gamit mankema si jafek	The ballerina barbecues the corn	ballerina, boxer, corn, t-shirt
13	Si pripas gumbuma si jafek	The director salts the corn	director, boxer, corn, trousers
14	Si gamit gumbuma si firel	The ballerina salts the mushroom	ballerina, acrobat, mushroom, t-shirt
15	Si badut gumbuma si sonis	The clown salts the sausage	clown, magician, sausage, shirt
16	Si gamit gumbuma si sonis	The ballerina salts the sausage	ballerina, magician, sausage, t-shirt
17	Si badut gumbuma si jafek	The clown salts the corn	clown, boxer, corn, shirt
18	Si pripas gumbuma si firel	The director salts the mushroom	director, acrobat, mushroom, trousers
19	Si daram kipoma si oblung	The acrobat dries the t-shirt	acrobat, director, t-shirt, sausage
20	Si towok kipoma si kemei	The magician dries the shirt	magician, clown, shirt, corn
21	Si metin kipomema si zelan	The boxer dries the trousers	boxer, ballerina, trousers, mushroom
22	Si towok kipoma si oblung	The magician dries the t-shirt	magician, ballerina, t-shirt, corn
23	Si metin kipoma si kemei	The boxer dries the shirt	boxer, director, shirt, mushroom
24	Si daram kipoma si zelan	The acrobat dries the trousers	acrobat, clown, trousers, sausage
25	Si metin felima si oblung	The boxer irons the t-shirt	boxer, clown, t-shirt, mushroom
26	Si daram felima si kemei	The acrobat irons the shirt	acrobat, ballerina, shirt, sausage
27	Si towok felima si zelan	The magician irons the trousers	magician, director, trousers, corn
28	Si daram felima si oblung	The acrobat irons the trousers	acrobat, director, trousers, sausage
29	Si towok felima si oblung	The magician irons the t-shirt	magician, clown, t-shirt, corn
30	Si metin felima si kemei	The boxer irons the shirt	boxer, ballerina, shirt, mushroom
31	Si towok tambama si zelan	The magician sews the trousers	magician, ballerina, trousers, corn
32	Si metin tambama si oblung	The boxer sews the t-shirt	boxer, director, t-shirt, mushroom
33	Si daram tambama si kemei	The acrobat sews the shirt	acrobat, clown, shirt, sausage
34	Si metin tambama si zelan	The boxer sews the trousers	boxer, clown, trousers, mushroom
35	Si daram tambama si oblung	The acrobat sews the t-shirt	acrobat, ballerina, t-shirt, sausage
36	Si towok tambama si kemei	The magician sews the shirt	magician, director, shirt, corn

Table A.9.: Vocabulary-Test Items Exp. 2

Item	Noun	Translation	Depicted Objects
1	sonis	sausage	sausage, corn, mushroom, clown
2	jafek	corn	corn, mushroom, shirt, trousers
3	firel	mushroom	mushroom, sausage, magician, ballerina
4	oblung	t-shirt	t-shirt, trousers, shirt, mushroom
5	zelan	trousers	trousers, shirt, sausage, corn
6	kemei	shirt	shirt, t-shirt, mushroom, acrobat
7	badut	clown	clown, acrobat, corn, sausage
8	gamit	ballerina	ballerina, magician, trousers, t-shirt
9	pripas	director	director, boxer, clown, shirt
10	metin	boxer	boxer, clown, shirt, t-shirt
11	towok	magician	magician, ballerina, mushroom, trousers
12	daram	acrobat	acrobat, director, corn, sausage

B. Experimental Items Experiments 3, 4, & 5

Figure B.1.: Example picture verb learning/testing, Exp. 3 (Exp. 4)
felimema
'iron'

Table B.1.: Verb-Learning Items Exp. 3

Item	Verb	Translation	Depicted Character	Depicted Object
1	bermamema	eat	man	prawn
2	bermamema	eat	woman	prawn
3	mankemema	barbecue	man	prawn
4	mankemema	barbecue	woman	prawn
5	tambamema	sew	man	dress
6	tambamema	sew	woman	dress
7	felimema	iron	man	dress
8	felimema	iron	woman	dress
9	gumbumema	take	man	dress
10	gumbumema	take	woman	dress
11	kipamema	pick up	man	dress
12	kipamema	pick up	woman	dress

Table B.2.: Verb-Testing Items Exp. 3

Item	Verb	Translation	Depicted Character	Depicted Object
1	bermamema	eat	man	prawn
2	bermamema	eat	woman	prawn
3	mankemema	barbecue	man	prawn
4	mankemema	barbecue	woman	prawn
5	tambamema	sew	man	dress
6	tambamema	sew	woman	dress
7	felimema	iron	man	dress
8	felimema	iron	woman	dress
9	gumbumema	take	man	dress
10	gumbumema	take	woman	dress
11	kipamema	pick up	man	dress
12	kipamema	pick up	woman	dress

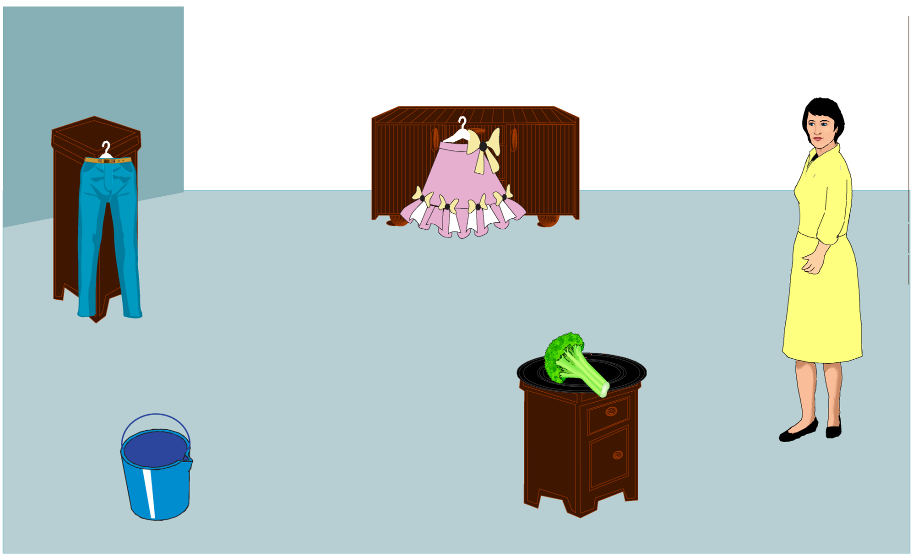


Figure B.2.: Example picture noun learning, Exp. 3, Condition NoRU
Si gadis mankemema si sonis
'The woman barbecues the broccoli'

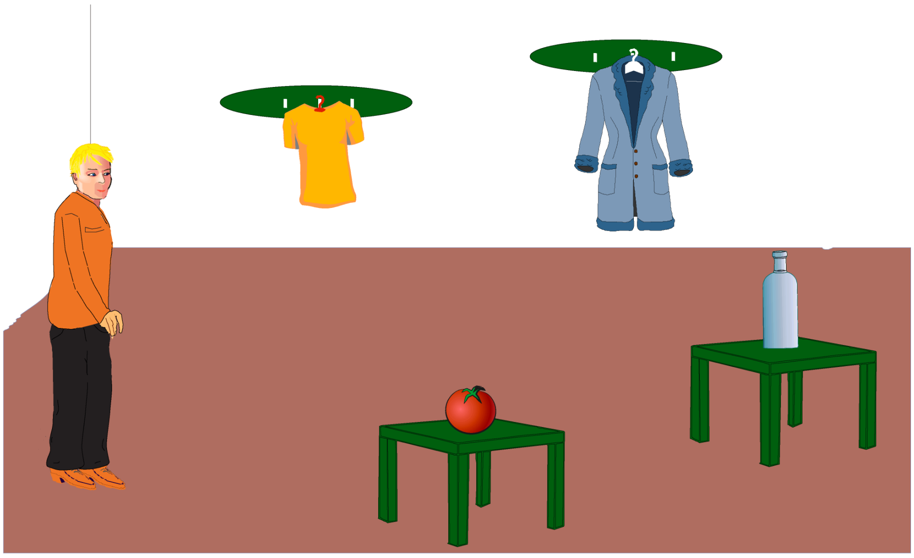


Figure B.3.: Example picture noun learning, Exp. 3, Condition LowRU Condition LowRU: *Si laki tambamema si kemei*, 'The man sews the t-shirt'

Figure B.4.: Example picture noun learning, Exp. 3, Condition HighRU *Si gadis kipamema si towok*, 'The woman takes the vase'

Table B.3.: Noun-learning sentences Exp. 3

Item	Sentence	Translation	Scene	Condition	Block
1	Si laki mankemema si sonis	The man barbecues the broccoli	broccoli, jacket, t-shirt, bucket	1	1
2	Si gadis mankemema si sonis	The woman barbecues the broccoli	broccoli, jacket, skirt, can	1	1
3	Si laki mankemema si sonis	The man barbecues the broccoli	broccoli, trousers, t-shirt, can	1	1
4	Si gadis mankemema si sonis	The woman barbecues the broccoli	broccoli, trousers, skirt, bucket	1	1
5	Si laki mankemema si jafek	The man barbecues the chicken	chicken, jacket, t-shirt, bucket	1	1
6	Si gadis mankemema si jafek	The woman barbecues the chicken	chicken, jacket, skirt, can	1	1
7	Si laki mankemema si jafek	The man barbecues the chicken	chicken, trousers, t-shirt, can	1	1
8	Si gadis mankemema si jafek	The woman barbecues the chicken	chicken, trousers, skirt, bucket	1	1
9	Si laki tambamema si kemei	The man sews the t-shirt	t-shirt, tomato, bottle, jacket	2	1
10	Si gadis tambamema si kemei	The woman sews the t-shirt	t-shirt, tomato, vase, trousers	2	1
11	Si laki tambamema si kemei	The man sews the t-shirt	t-shirt, toast , bottle, trousers	2	1
12	Si gadis tambamema si kemei	The woman sews the t-shirt	t-shirt, toast, vase, jacket	2	1
13	Si laki tambamema si gamit	The man sews the skirt	skirt, tomato, bottle jacket	2	1
14	Si gadis tambamema si gamit	The woman sews the skirt	skirt, tomato, vase, trousers	2	1
15	Si laki tambamema si gamit	The man sews the skirt	skirt, toast, bottle, trousers	2	1
16	Si gadis tambamema si gamit	The woman sews the skirt	skirt, toast, vase, jacket	2	1
17	Si laki kipamema si badut	The man takes the bottle	bottle, bucket, jacket, tomato	3	1
18	Si gadis kipamema si badut	The woman takes the bottle	bottle, bucket, trousers, toast	3	1
19	Si laki kipamema si badut	The man takes the bottle	bottle, can, jacket, toast	3	1
20	Si gadis kipamema si badut	The woman takes the bottle	bottle, can, trousers, tomato	3	1
21	Si laki kipamema si towok	The man takes the vase	vase, bucket, jacket, tomato	3	1
22	Si gadis kipamema si towok	The woman takes the vase	vase, bucket, trousers, toast	3	1
23	Si laki kipamema si towok	The man takes the vase	vase, can, jacket, toast	3	1
24	Si gadis kipamema si towok	The woman takes the vase	vase, can, trousers tomato	3	1
25	Si laki felimema si oblung	The man irons the jacket	jacket, broccoli, tomato, bottle	1	2
26	Si gadis felimema si oblung	The woman irons the jacket	jacket, broccoli, toast, vase	1	2
27	Si laki felimema si oblung	The man irons the jacket	jacket, chicken, tomato, vase	1	2
28	Si gadis felimema si oblung	The woman irons the jacket	jacket, chicken, toast, bottle	1	2
29	Si laki felimema si zelan	The man irons the trousers	trousers, broccoli, tomato, bottle	1	2
30	Si gadis felimema si zelan	The woman irons the trousers	trousers, broccoli, toast, vase	1	2
31	Si laki felimema si zelan	The man irons the trousers	trousers, chicken, tomato, vase	1	2
32	Si gadis felimema si zelan	The woman irons the trousers	trousers, chicken, toast , bottle	1	2
33	Si laki bermamema si pripas	The man eats the tomato	tomato, t-shirt, bucket, broccoli	2	2
34	Si gadis bermamema si pripas	The woman eats the tomato	tomato, t-shirt, can, chicken	2	2
35	Si laki bermamema si pripas	The man eats the tomato	tomato, skirt, bucket, chicken	2	2
36	Si gadis bermamema si pripas	The woman eats the tomato	tomato, skirt, can, broccoli	2	2
37	Si laki bermamema si firel	The man eats the toast	toast, t-shirt, bucket, broccoli	2	2
38	Si gadis bermamema si firel	The woman eats the toast	toast, t-shirt, can, chicken	2	2
39	Si laki bermamema si firel	The man eats the toast	toast, skirt, bucket, chicken	2	2
40	Si gadis bermamema si firel	The woman eats the toast	toast, skirt, can, broccoli	2	2
41	Si laki gumbumema Si metin	The man points at the bucket	bucket, bottle, broccoli, t-shirt	3	2
42	Si gadis gumbumema Si metin	The woman points at the bucket	bucket, bottle, chicken, skirt	3	2
43	Si laki gumbumema Si metin	The man points at the bucket	bucket, vase, broccoli, skirt	3	2
44	Si gadis gumbumema Si metin	The woman points at the bucket	bucket, vase, chicken, t-shirt	3	2
45	Si laki gumbumema si daram	The man points at the watering can	watering can, bottle, broccoli, t-shirt	3	2
46	Si gadis gumbumema si daram	The woman points at the watering can	watering can, bottle, chicken, skirt	3	2
47	Si laki gumbumema si daram	The man points at the watering can	watering can, vase, broccoli, skirt	3	2
48	Si gadis gumbumema si daram	The woman points at the watering can	watering can, vase, chicken, t-shirt	3	2

Figure B.5.: Example picture noun testing, Exp. 3
si gadis
'the woman'

Table B.4.: Vocabulary-Test Items Exp. 3

Item	Noun	Translation	Depicted Objects
1	sonis	broccoli	chicken, t-shirt, skirt, man1, woman1
2	jafek	chicken	broccoli, skirt, t-shirt, woman1, man1
3	kemei	t-shirt	skirt, bottle, vase, man2, woman1
4	gamit	skirt	t-shirt, vase, bottle, woman2, man1
5	badut	bottle	vase, broccoli, chicken, man2, woman1
6	towok	vase	bottle, chicken, broccoli, woman2, man1
7	laki	man	woman, t-shirt, skirt, broccoli, chicken
8	oblung	jacket	trousers, tomato, toast, man2, woman2
9	zelan	trousers	jacket, toast, tomato, woman2, man2
10	pripas	tomato	toast, bucket, can, man1, woman2
11	firel	toast	tomato, can, bucket, woman1, man2
12	metin	bucket	can, jacket, trousers, man1, woman2
13	daram	can	bucket, trousers, jacket, woman1, man2
14	gadis	woman	man, jacket, trousers, bucket, can

Table B.5.: Verb-Learning Items Exp. 4 & 5

Item	Verb	Translation	Depicted Character	Depicted Object
1	bermamema	eat	man	pepper
2	bermamema	eat	woman	sausage
3	mankemema	sew	man	scarf
4	mankemema	sew	woman	shirt
5	tambamema	take	man	vase
6	tambamema	take	woman	pepper
7	gumbumema	point at	man	vase
8	gumbumema	point at	woman	pepper

Table B.6.: Verb-Testing Items Exp. 4 & 5

Item	Verb	Translation	Depicted Character	Depicted Object
1	bermamema	eat	man	prawn
2	bermamema	eat	woman	prawm
3	mankemema	sew	man	dress
4	mankemema	sew	woman	dress
5	tambamema	take	man	dress
6	tambamema	take	woman	dress
7	gumbumema	point at	man	dress
8	gumbumema	point at	woman	dress

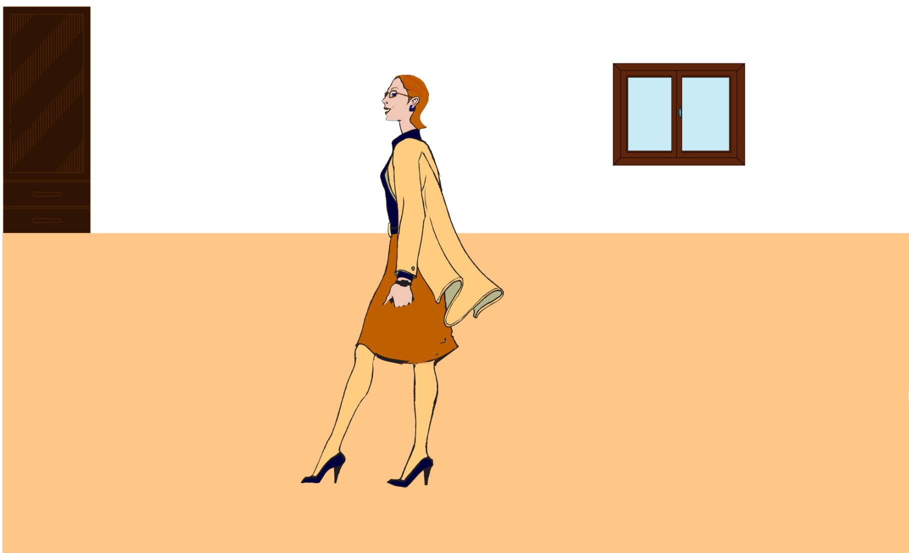


Figure B.6.: Example picture noun learning, Exp. 4, Condition N, Trial Type 1
Si gades tambamema si towok
'The woman takes the cucumber/jacket'

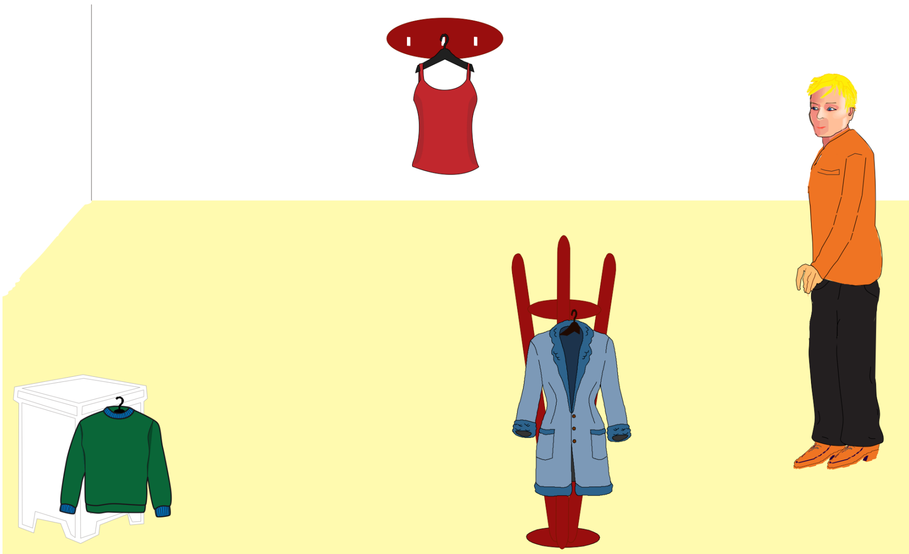


Figure B.7.: Example picture noun learning, Exp. 4, Condition N, Trial Type 2
Si laki gumbumema si towok
'The man points at the cucumber/jacket'

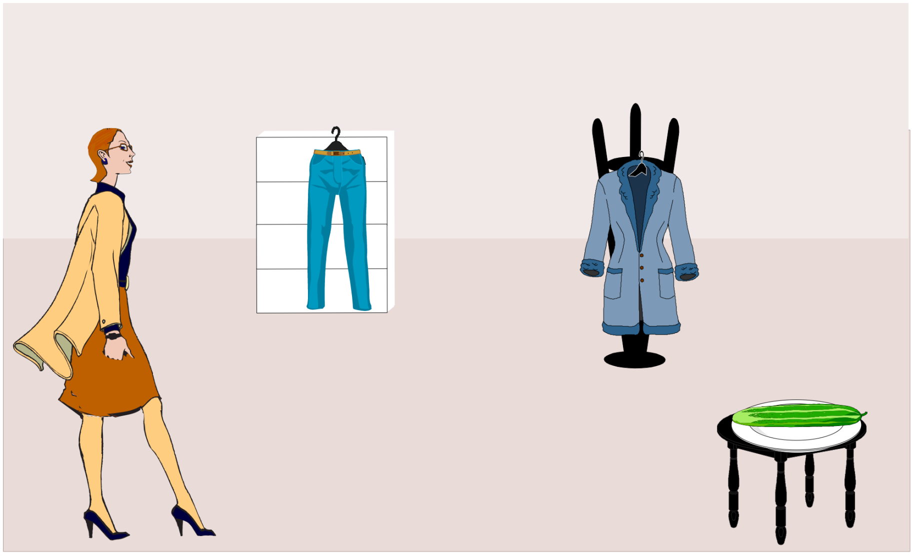


Figure B.8.: Example picture noun learning, Exp. 4, Condition N, Trial Type 3
Si gades tambamema si towok
'The woman takes the cucumber/jacket'

Figure B.9.: Example picture noun learning, Exp. 4, Condition N, Trial Type 4
Si laki gumbumema si towok
'The man points at the cucumber/jacket'

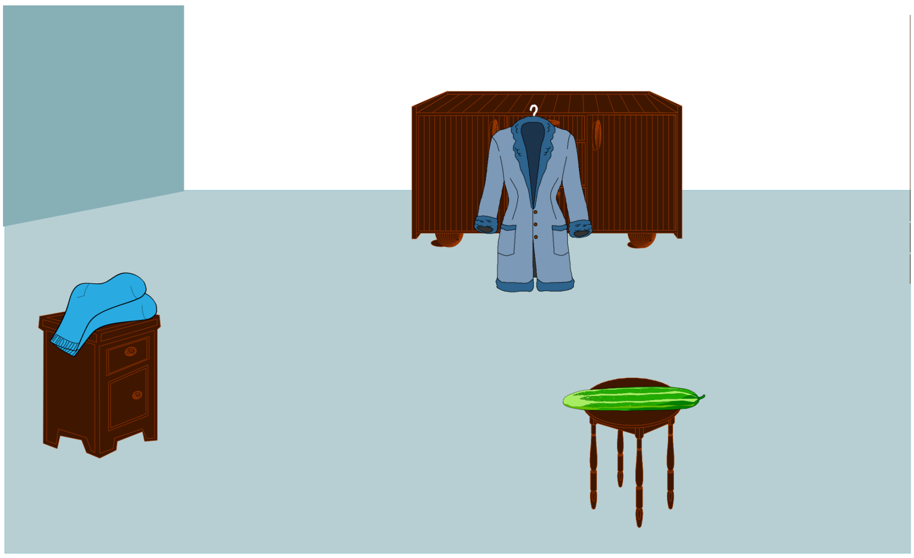


Figure B.10.: Example picture noun learning, Exp. 4, Condition N, Trial Type 5
Si gades tambamema si towok
'The woman takes the cucumber/jacket'

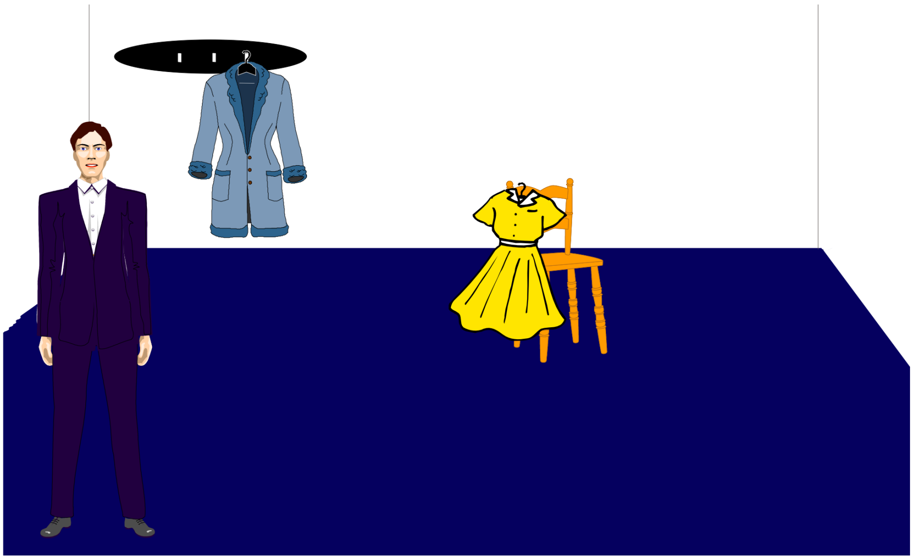


Figure B.11.: Example picture noun learning, Exp. 4, Condition N, Trial Type 6
Si laki gumbumema si towok
'The man points at the cucumber/jacket'

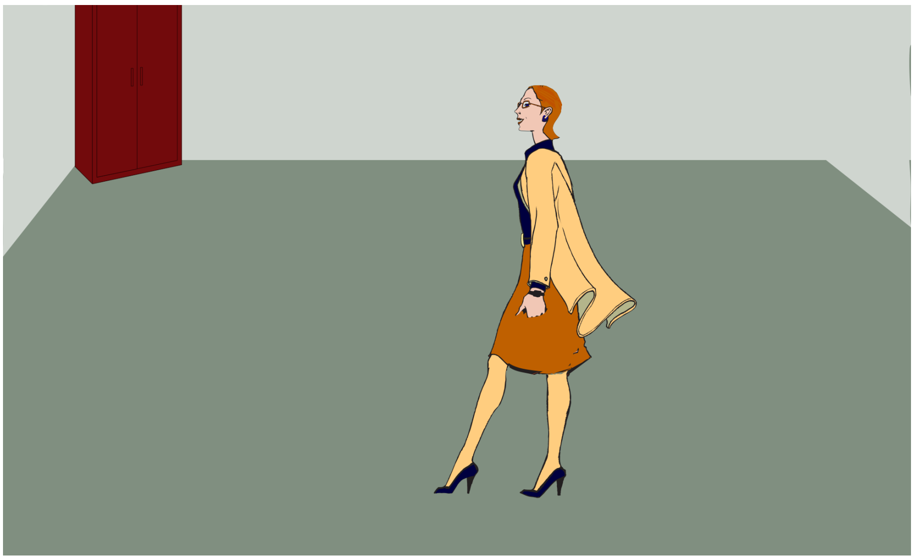


Figure B.12.: Example picture noun learning, Exp. 4, Condition R, Trial Type 1
Si gades mankemema si oblung
'The woman sews the t-shirt/fries'

Figure B.13.: Example picture noun learning, Exp. 4, Condition R, Trial Type 2
Si laki mankemema si oblung
'The man sews the t-shirt/fries'

Figure B.14.: Example picture noun learning, Exp. 4, Condition R, Trial Type 3
Si gades tambamema si oblung
'The woman takes the t-shirt/fries'

Figure B.15.: Example picture noun learning, Exp. 4, Condition R, Trial Type 4
Si laki gumbumema si oblung
'The man points at the t-shirt/fries'

Figure B.16.: Example picture noun learning, Exp. 4, Condition R, Trial Type 5
Si gades tambamema si oblung
'The woman takes the t-shirt/fries'

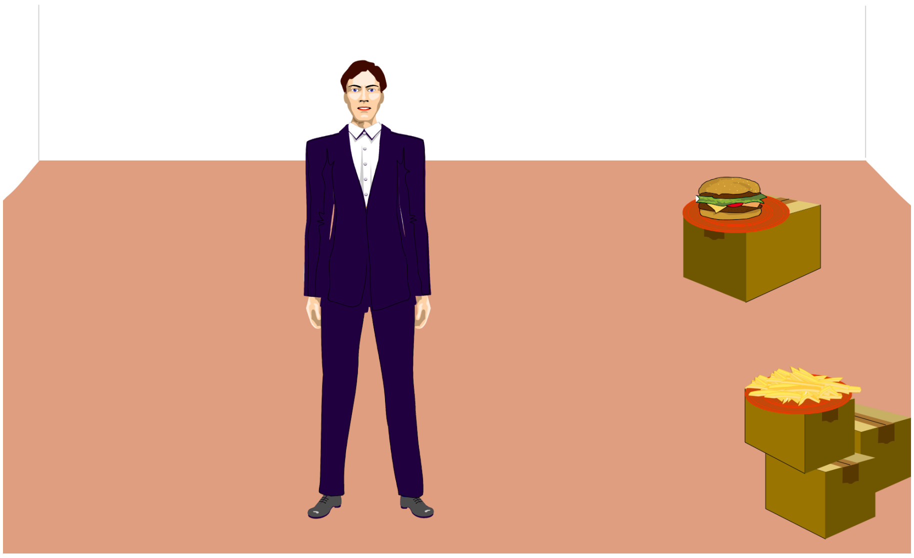


Figure B.17.: Example picture noun learning, Exp. 4, Condition R, Trial Type 6
Si laki makemema si oblung
'The man sews the t-shirt/fries'

Table B.7.: Noun-learning sentences Exp. 4

Item	Sentence	Translation	Scene	Condition
1	Si gades tambamema si towok.	The woman takes the cucumber/jacket	girl	Non-restrictive
2	Si laki gumbumema si towok.	The man points at the cucumber/jacket	boy top jacket jumper	Non-restrictive
3	Si gades tambamema si towok.	The woman takes the cucumber/jacket	girl cucumber jacket jeans	Non-restrictive
4	Si laki gumbumema si towok.	The man points at the cucumber/jacket	cucumber jacket pizza	Non-restrictive
5	Si gades tambamema si towok.	The woman takes the cucumber/jacket	cucumber jacket socks	Non-restrictive
6	Si laki gumbumema si towok.	The man points at the cucumber/jacket	boy jacket dress	Non-restrictive
7	Si gades mankemema si oblung.	The woman sews the t-shirt/fries	girl	Restrictive
8	Si laki mankemema si oblung.	The man sews the t-shirt/fries	boy banana fries sausage	Restrictive
9	Si gades tambamema si oblung.	The woman takes the t-shirt/fries	girl tshirt fries melon	Restrictive
10	Si laki gumbumema si oblung.	The man points at the t-shirt/fries	tshirt fries jacket	Restrictive
11	Si gades tambamema si oblung.	The woman takes the t-shirt/fries	tshirt fries cucumber	Restrictive
12	Si laki makemema si oblung.	The man sews the t-shirt/fries	boy fries hamburger	Restrictive
13	Si laki bermamema si sonis.	The man eats the banana/socks	boy	Restrictive
14	Si gades bermamema si sonis.	The woman eats the banana/socks	girl jeans socks hat	Restrictive
15	Si laki tambamema si sonis.	The man takes the banana/socks	boy banana socks top	Restrictive
16	Si gades gumbumema si sonis.	The woman points at the banana/socks	banana socks sausage	Restrictive
17	Si laki tambamema si sonis.	The man takes thebanana/socks	banana socks shorts	Restrictive
18	Si gades bermamema si sonis.	The woman eats the banana/socks	girl socks skirt	Restrictive
19	Si laki bermamema si kemei.	The woman eats the pizza/shorts	boy	Restrictive
20	Si gades bermamema si kemei.	The man eats the pizza/shorts	girl tshirt shorts jumper	Restrictive
21	Si laki tambamema si kemei.	The woman takes the pizza/shorts	boy pizza shorts dress	Restrictive
22	Si gades gumbumema si kemei.	The man points at the pizza/shorts	pizza shorts apple	Restrictive
23	Si laki tambamema si kemei.	The woman takes the pizza/shorts	pizza shorts skirt	Restrictive
24	Si gades bermamema si kemei.	The man eats the pizza/shorts	girl shorts socks	Restrictive
25	Si gades tambamema si gamit.	The woman takes the carrot/dress	girl	Restrictive
26	Si laki gumbumema si gamit.	The man takes the carrot/dress	boy dress top jeans	Non-restrictive
27	Si gades tambamema si gamit.	The woman takes the carrot/dress	girl carrot top hat	Non-restrictive
28	Si laki gumbumema si gamit.	The man takes the carrot/dress	carrot top apple	Non-restrictive
29	Si gades tambamema si gamit.	The woman takes the carrot/dress	carrot top socks	Non-restrictive
30	Si laki gumbumema si gamit.	The man takes the carrot/dress	boy top shorts	Non-restrictive
31	Si laki tambamema si badut.	The man takes the corn/jumper	boy	Non-restrictive
32	Si gades gumbumema si badut.	The woman takes the corn/jumper	girl hat jumper skirt	Non-restrictive
33	Si laki tambamema si badut.	The man takes the corn/jumper	boy corn jumper shorts	Non-restrictive
34	Si gades gumbumema si badut.	The woman takes the corn/jumper	corn jumper banana	Non-restrictive
35	Si laki tambamema si badut.	The man takes the corn/jumper	corn jumper tshirt	Non-restrictive
36	Si gades gumbumema si badut.	The woman takes the corn/jumper	girl jumper top	Non-restrictive
37	Si laki mankemema si zelan.	The man sews the skirt/melon	boy	Restrictive
38	Si gades mankemema si zelan.	The woman sews the skirt/melon	girl hamburger melon corn	Restrictive
39	Si laki tambamema si zelan.	The man takes the skirt/melon	boy skirt melon fries	Restrictive
40	Si gades gumbumema si zelan.	The woman points at the skirt/melon	skirt melon jacket	Restrictive
41	Si laki tambamema si zelan.	The man takes the skirt/melon	skirt melon carrot	Restrictive
42	Si gades makemema si zelan.	The woman sews the skirt/melon	girl melon banana	Restrictive
43	Si gades mankemema si pripas.	The woman sews the jeans/sausage	girl	Restrictive
44	Si laki mankemema si pripas.	The man sews the jeans/sausage	boy cucumber sausage apple	Restrictive
45	Si gades tambamema si pripas.	The woman takes the jeans/sausage	girl jeans sausage hamburger	Restrictive
46	Si laki gumbumema si pripas.	The man points at the jeans/sausage	jeans sausage tshirt	Restrictive
47	Si gades tambamema si pripas.	The woman takes the jeans/sausage	jeans sausage pizza	Restrictive
48	Si laki makemema si pripas.	The man sews the jeans/sausage	boy sausage melon	Restrictive
49	Si gades tambamema si metin.	The woman takes the dress/apple	girl	Non-restrictive
50	Si laki gumbumema si metin.	The man takes the dress/apple	boy carrot apple fries	Non-restrictive
51	Si gades tambamema si metin.	The woman takes the dress/apple	girl dress apple melon	Non-restrictive
52	Si laki gumbumema si metin.	The man takes the dress/apple	dress apple jumper	Non-restrictive
53	Si gades tambamema si metin.	The woman takes the dress/apple	dress apple corn	Non-restrictive
54	Si laki gumbumema si metin.	The man takes the dress/apple	boy apple sausage	Non-restrictive
55	Si gades tambamema si daram.	The woman takes the hat/hamurger	girl	Non-restrictive
56	Si laki gumbumema si daram.	The man takes the hat/hamurger	boy corn hamburger cucumber	Non-restrictive
57	Si gades tambamema si daram.	The woman takes the hat/hamurger	girl hat hamburger pizza	Non-restrictive
58	Si laki gumbumema si daram.	The man takes the hat/hamurger	hat hamburger jacket	Non-restrictive
59	Si gades tambamema si daram.	The woman takes the hat/hamurger	hat hamburger fries	Non-restrictive
60	Si laki gumbumema si daram.	The man takes the hat/hamurger	girl hamburger	Non-restrictive

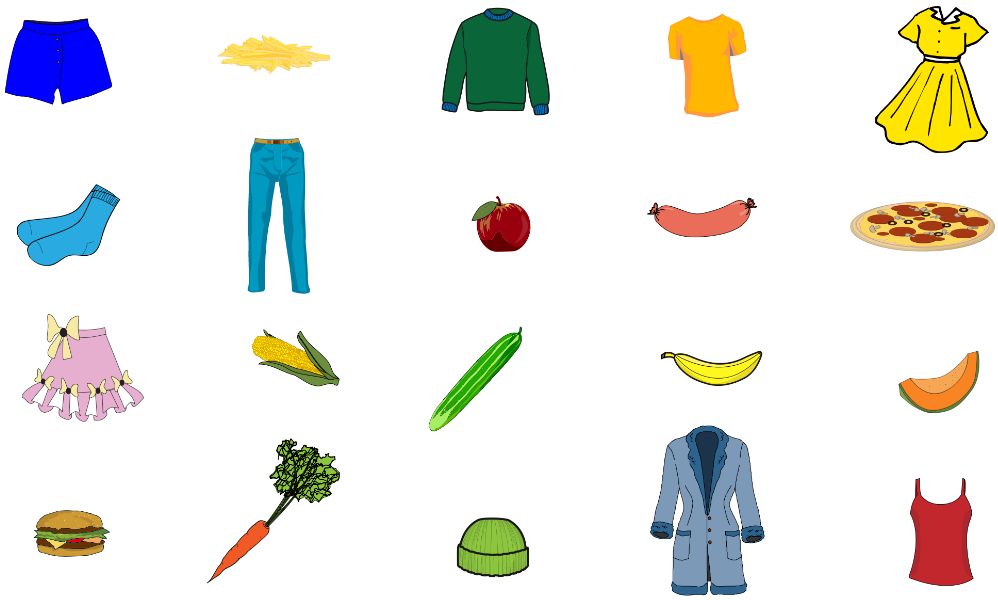


Figure B.18.: Example picture noun testing, Exp. 4
si towok
'the cucumber/jacket'

Table B.8.: Vocabulary-Test Items Exp. 4

Item	Noun	Translation	Depicted Objects
1	towok	cucumber/jacket	all
2	oblung	t-shirt/fries	all
3	sonis	banana/socks	all
4	kemei	pizza/shorts	all
5	gamit	carrot/top	all
6	badut	corn/jumper	all
7	zelan	skirt/melon	all
8	pripas	jeans/sausage	all
9	metin	dress/apple	all
10	daram	hat/hamburger	all

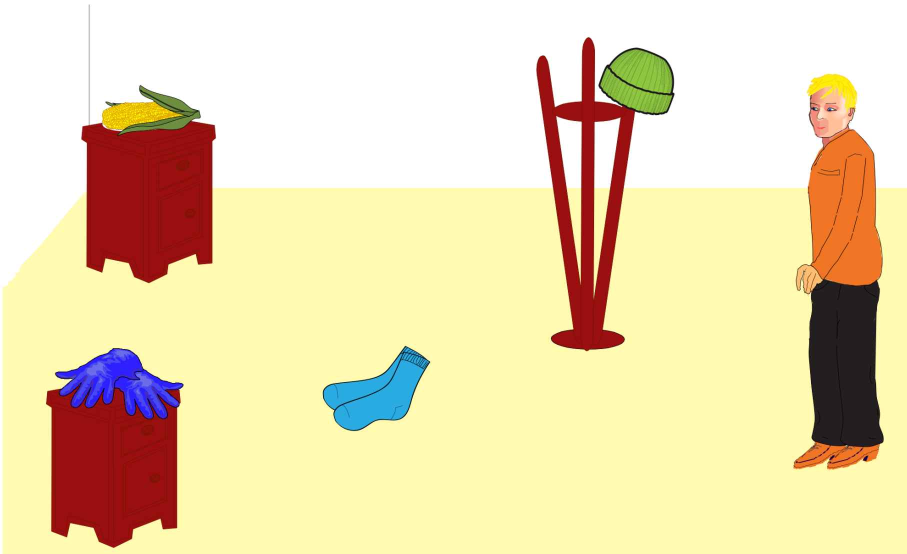


Figure B.19.: Example picture noun learning, Exp. 5, Condition N, Trial 1
Si laki bermamema si sonis
'The man eats the SONIS'

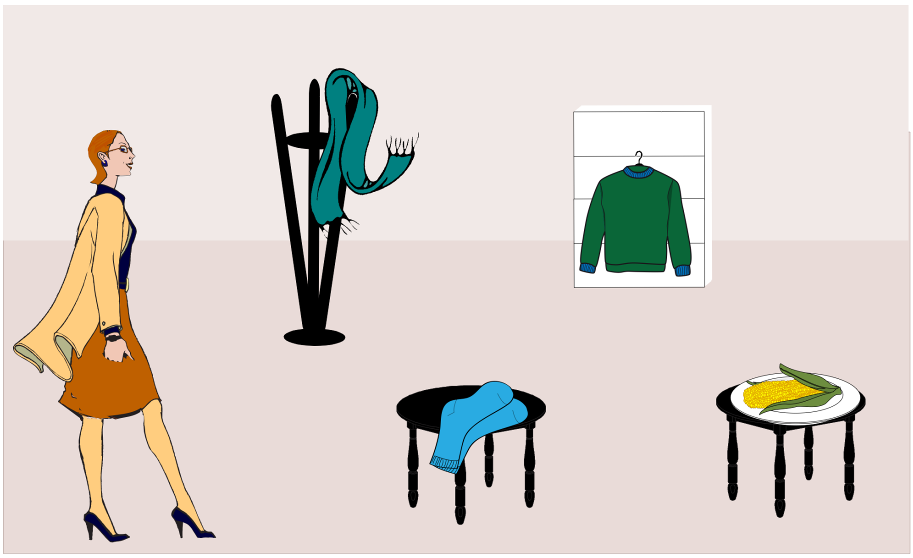


Figure B.20.: Example picture noun learning, Exp. 5, Condition N, Trial 2
Si gadis bermamema si sonis
'The woman eats the SONIS'

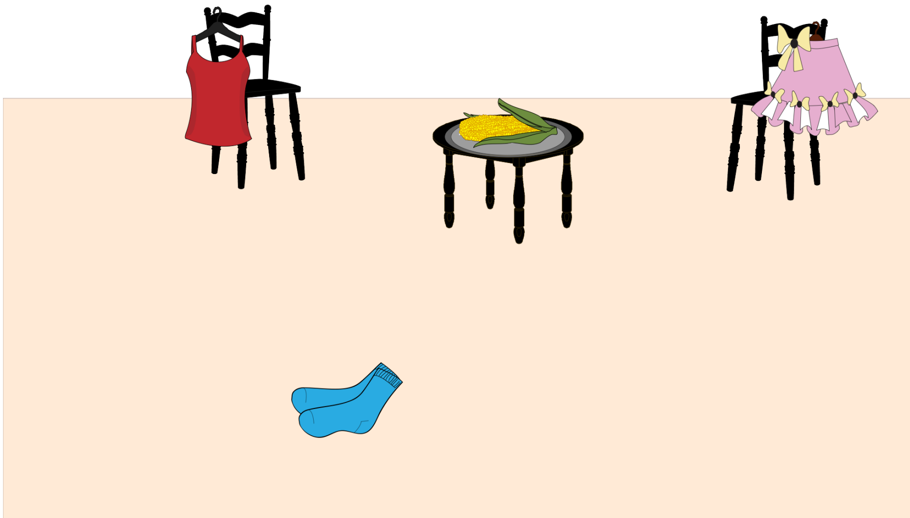


Figure B.21.: Example picture noun learning, Exp. 5, Condition N, Trial 3
Si gadis bermamema si sonis
'The woman eats the SONIS'

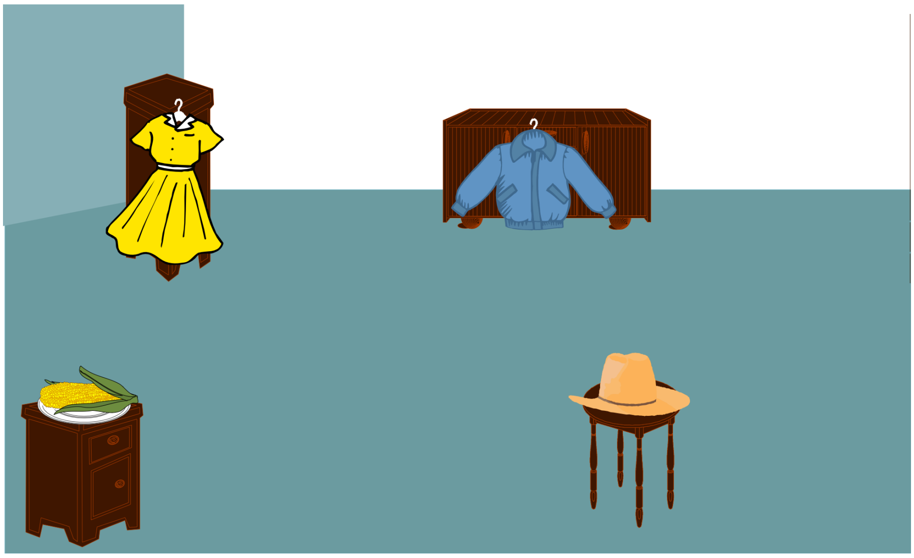


Figure B.22.: Example picture noun learning, Exp. 5, Condition N, Trial 4
Si gadis bermamema si sonis
'The woman eats the SONIS'

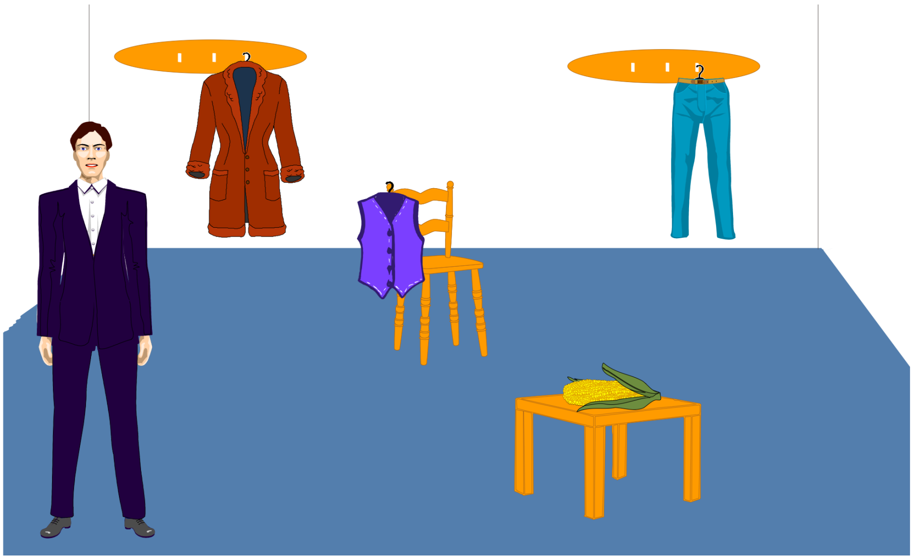


Figure B.23.: Example picture noun learning, Exp. 5, Condition N, Trial 5
Si laki bermamema si sonis
'The man eats the SONIS'

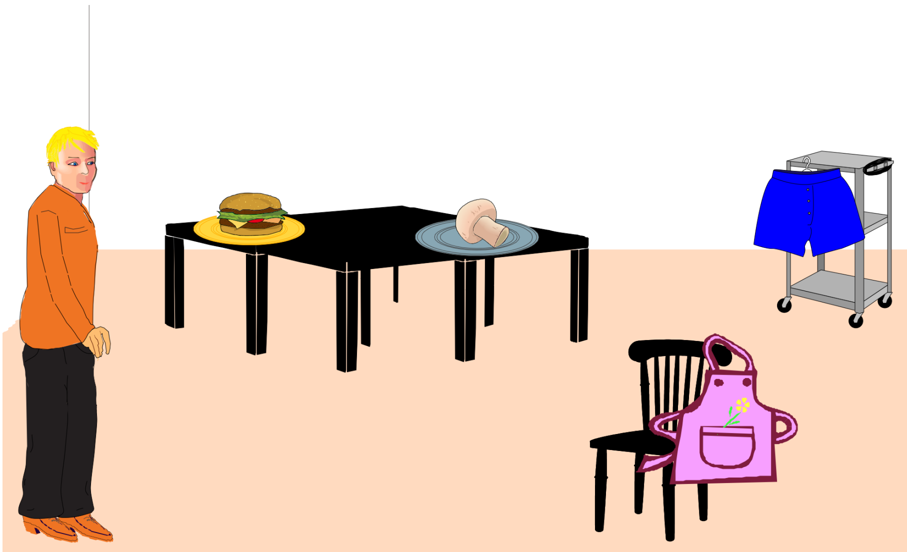


Figure B.24.: Example picture noun learning, Exp. 5, Condition N, Trial 6
Si laki tambamema si sonis
'The man takes the SONIS'

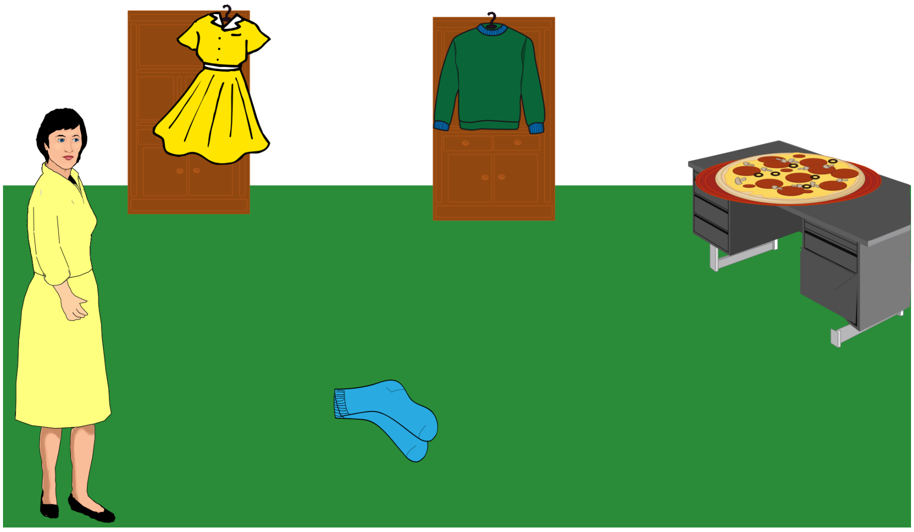


Figure B.25.: Example picture noun learning, Exp. 5, Condition R, Trial 1
Si gadis gumbumema si firel
'The woman points at the FIREL'

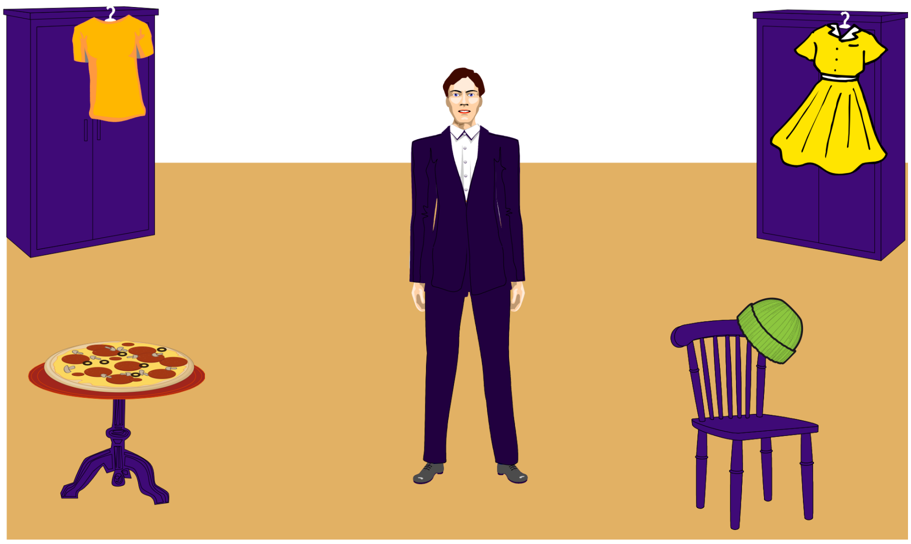


Figure B.26.: Example picture noun learning, Exp. 5, Condition R, Trial 2
Si laki tambamema si firel
'The man takes the FIREL'

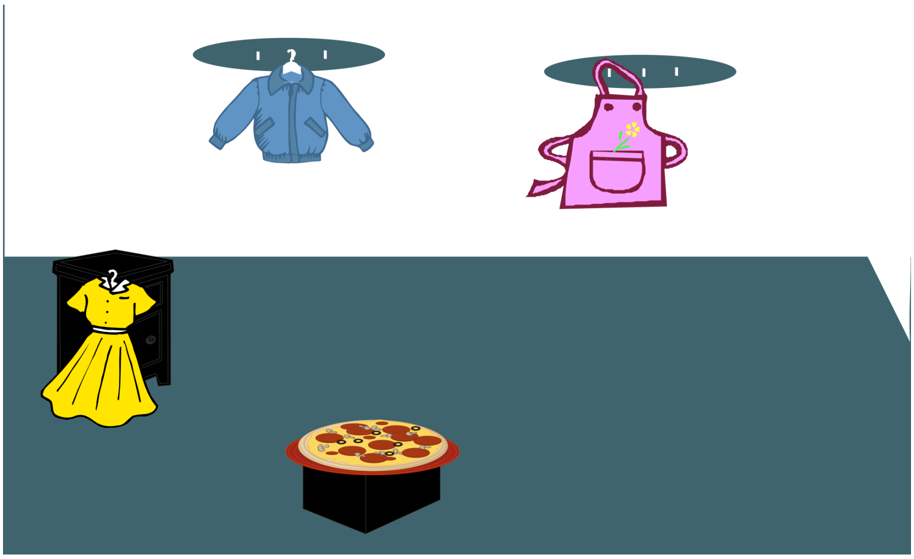


Figure B.27.: Example picture noun learning, Exp. 5, Condition R, Trial 3
Si gadis gumbumema si firel
'The woman points at the FIREL'

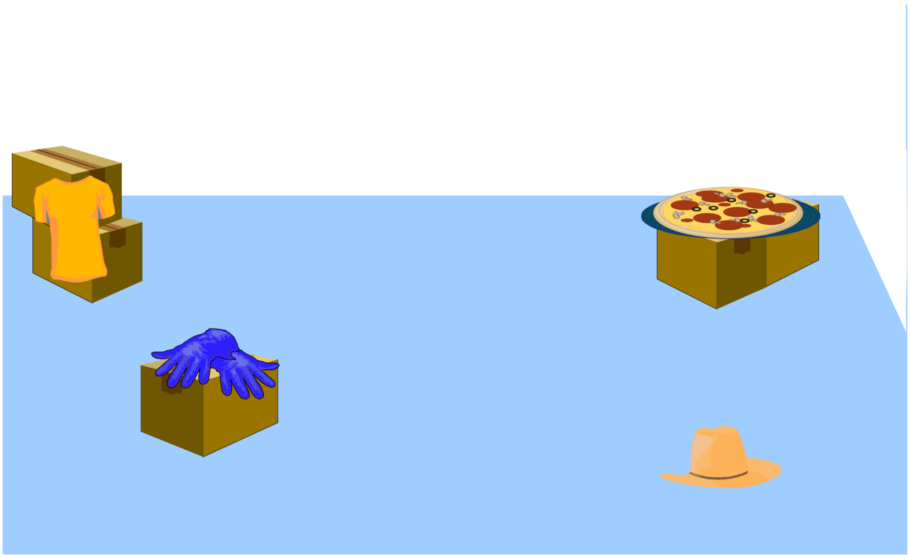


Figure B.28.: Example picture noun learning, Exp. 5, Condition R, Trial 4
Si laki tambamema si firel
'The man points at the FIREL'

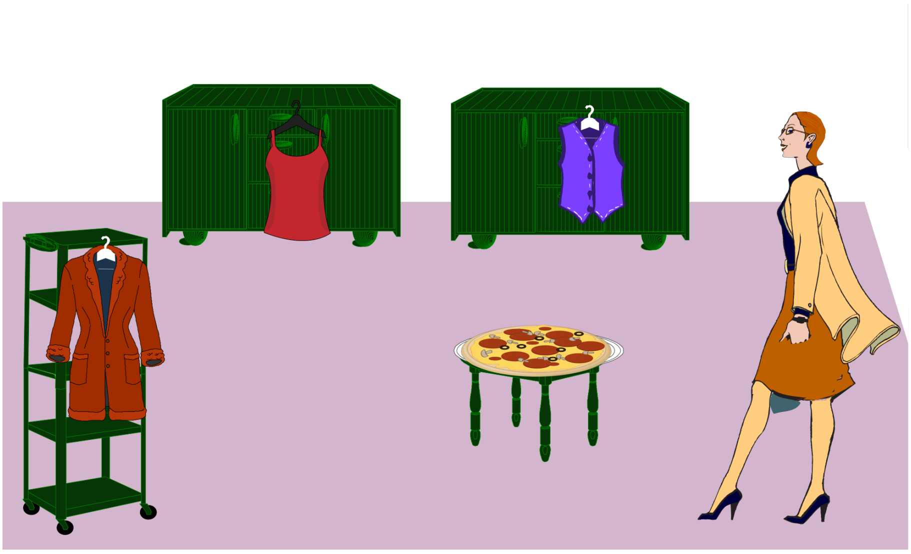


Figure B.29.: Example picture noun learning, Exp. 5, Condition R, Trial 5
Si laki gumbumema si firel
'The man points at the FIREL'

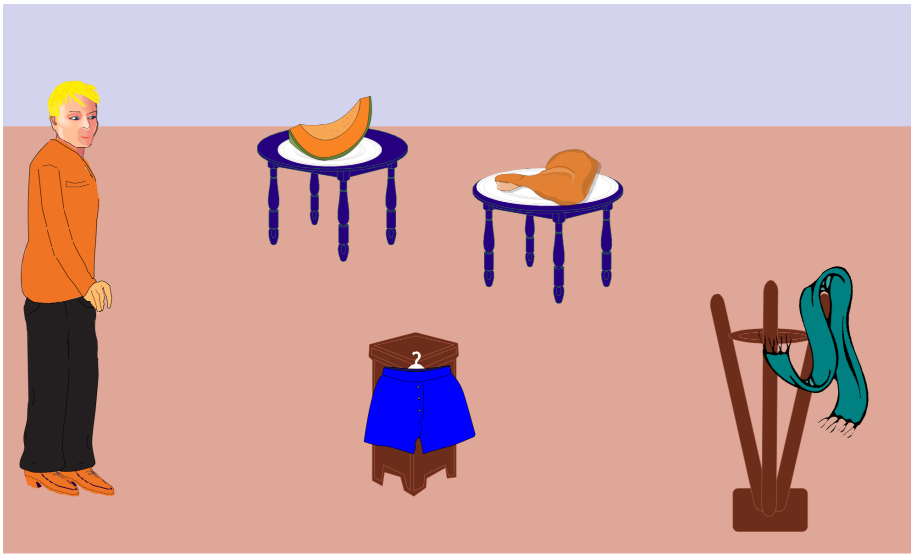


Figure B.30.: Example picture noun learning, Exp. 5, Condition R, Trial 6
Si laki tambamema si firel
'The man takes the FIREL'

Table B.9.: Noun-learning sentences Exp. 5

Item	Sentence	Translation	Scene	Condition
1	Si laki bermamema si sonis.	The man eats the SONIS.	corn socks gloves hat man	restrictive
2	Si gadis bermamema si sonis.	The woman eats the SONIS.	corn socks jumper scarf woman	restrictive
3	Si gadis bermamema si sonis.	The woman eats the SONIS.	corn socks top skirt	restrictive
4	Si gadis bermamema si sonis.	The woman eats the SONIS.	corn cap jacket dress	restrictive
5	Si laki bermamema si sonis.	The man eats the SONIS.	corn vest jeans coat man	restrictive
6	Si laki tambamema si sonis.	The man takes the SONIS.	hamburger mushroom shorts apron man	restrictive
7	Si gadis bermamema si badut.	The woman eats the BADUT.	apple hat shirt socks woman	restrictive
8	Si laki bermamema si badut.	The man eats the BADUT.	apple hat gloves scarf man	restrictive
9	Si laki bermamema si badut.	The man eats the BADUT.	apple hat top dress	restrictive
10	Si gadis bermamema si badut.	The woman eats the BADUT.	apple cap jacket coat	restrictive
11	Si gadis bermamema si badut.	The woman eats the BADUT.	apple vest jeans apron woman	restrictive
12	Si laki tambamema si badut.	The man takes the BADUT.	fries ananas shorts skirt man	restrictive
13	Si gadis bermamema si kemei.	The woman eats the KEMEI.	cucumber scarf jeans socks woman	restrictive
14	Si laki bermamema si kemei.	The man eats the KEMEI.	cucumber scarf jumper hat man	restrictive
15	Si laki bermamema si kemei.	The man eats the KEMEI.	cucumber scarf jacket skirt	restrictive
16	Si laki bermamema si kemei.	The man eats the KEMEI.	cucumber cap gloves coat	restrictive
17	Si gadis bermamema si kemei.	The woman eats the KEMEI.	cucumber vest shirt apron woman	restrictive
18	Si gadis tambamema si kemei.	The woman takes the KEMEI.	sausage cheese shorts dress woman	restrictive
19	Si gadis bermamema si towok.	The woman eats the TOWOK.	banana skirt gloves socks woman	restrictive
20	Si laki bermamema si towok.	The man eats the TOWOK.	banana skirt jeans hat man	restrictive
21	Si gadis bermamema si towok.	The woman eats the TOWOK.	banana skirt jumper dress	restrictive
22	Si gadis bermamema si towok.	The woman eats the TOWOK.	banana cap shorts scarf	restrictive
23	Si laki bermamema si towok.	The man eats the TOWOK.	banana vest top apron man	restrictive
24	Si laki tambamema si towok.	The man takes the TOWOK.	melon chicken shirt coat man	restrictive
25	Si gadis gumbumema si firel.	The woman points at the FIREL.	pizza dress jumper socks woman	Non-restrictive
26	Si laki tambamema si firel.	The man takes the FIREL.	pizza dress shirt hat man	Non-restrictive
27	Si gadis gumbumema si firel.	The woman points at the FIREL.	pizza dress jacket apron	Non-restrictive
28	Si laki tambamema si firel.	The man takes the FIREL.	pizza cap gloves skirt	Non-restrictive
29	Si laki gumbumema si firel.	The man points at the FIREL.	pizza vest top coat woman	Non-restrictive
30	Si laki tambamema si firel.	The man takes the FIREL.	chicken melon shorts scarf man	Non-restrictive
31	Si gadis tambamema si oblung.	The woman takes the OBLUNG.	carrot coat shirt dress woman	Non-restrictive
32	Si laki gumbumema si oblung.	The man points at OBLUNG.	carrot coat jumper hat man	Non-restrictive
33	Si laki tambamema si oblung.	The man takes the OBLUNG.	carrot coat top scarf	Non-restrictive
34	Si laki gumbumema si oblung.	The man points at the OBLUNG.	carrot cap jacket skirt	Non-restrictive
35	Si gadis tambamema si oblung.	The woman takes the OBLUNG.	carrot vest shorts apron woman	Non-restrictive
36	Si gadis gumbumema si oblung.	The woman points at the OBLUNG.	cheese sausage jeans socks woman	Non-restrictive
37	Si laki tambamema si zelan.	The man takes the ZELAN.	bread apron jumper skirt man	Non-restrictive
38	Si laki gumbumema si zelan.	The man points at ZELAN.	bread apron shirt coat man	Non-restrictive
39	Si gadis tambamema si zelan.	The woman takes the ZELAN.	bread apron top scarf	Non-restrictive
40	Si gadis gumbumema si zelan.	The woman points at the ZELAN.	bread cap jeans socks	Non-restrictive
41	Si gadis tambamema si zelan.	The woman takes the ZELAN.	bread hat jacket dress woman	Non-restrictive
42	Si laki gumbumema si zelan.	The man points at the ZELAN.	ananas fries shorts vest man	Non-restrictive
43	Si gadis tambamema si pripas.	The woman takes the PRIPAS.	cake vest shirt scarf woman	Non-restrictive
44	Si laki gumbumema si pripas.	The man points at PRIPAS.	cake vest jacket socks man	Non-restrictive
45	Si laki tambamema si pripas.	The man takes the PRIPAS.	cake vest jeans skirt	Non-restrictive
46	Si gadis gumbumema si pripas.	The woman points at the PRIPAS.	cake gloves jumper dress	Non-restrictive
47	Si laki tambamema si pripas.	The man takes the PRIPAS.	cake apron top coat man	Non-restrictive
48	Si gadis gumbumema si pripas.	The woman points at the PRIPAS.	mushroom hamburger gloves hat woman	Non-restrictive
49	Si laki mankemema si metin.	The man sews the METIN.	shirt hamburger cake fries man	restrictive
50	Si gadis mankemema si metin.	The woman sews the METIN.	shirt hamburger apple sausage woman	restrictive
51	Si laki mankemema si metin.	The man sews the METIN.	shirt hamburger cucumber melon	restrictive
52	Si gadis mankemema si metin.	The woman sews the METIN.	shirt bread banana chicken	restrictive
53	Si gadis mankemema si metin.	The woman sews the METIN.	shirt mushroom pizza cheese woman	restrictive
54	Si laki gumbumema si metin.	The man points at the METIN.	socks vest carrot ananas man	restrictive
55	Si gadis mankemema si daram.	The woman sews the DARAM.	jumper fries cake hamburger woman	restrictive
56	Si laki mankemema si daram.	The man sews the DARAM.	jumper fries corn sausage man	restrictive
57	Si gadis mankemema si daram.	The woman sews the DARAM.	jumper fries banana melon	restrictive
58	Si gadis mankemema si daram.	The woman sews the DARAM.	jumper bread cucumber cheese	restrictive
59	Si laki mankemema si daram.	The man sews the DARAM.	jumper mushroom carrot ananas man	restrictive
60	Si laki gumbumema si daram.	The man points at the DARAM.	hat apron pizza chicken man	restrictive

Table B.10.: Noun-learning sentences Exp. 5 cont.

Item	Sentence	Translation	Scene	Condition
60	Si laki gumbumema si daram.	The man points at the DARAM.	hat apron pizza chicken man	Restrictive
61	Si laki mankemema si kidir.	The man sews the KIDIR.	top sausage cake melon man	restrictive
62	Si gadis mankemema si kidir.	The woman sews the KIDIR.	top sausage banana fries woman	restrictive
63	Si gadis mankemema si kidir.	The woman sews the KIDIR.	top sausage apple hamburger	restrictive
64	Si gadis mankemema si kidir.	The woman sews the KIDIR.	top bread carrot ananas	restrictive
65	Si laki mankemema si kidir.	The man sews the KIDIR.	top mushroom corn chicken man	restrictive
66	Si laki gumbumema si kidir.	The man points at the KIDIR.	scarf coat pizza cheese man	restrictive
67	Si gadis mankemema si bintang.	The woman sews the BINTANG.	jacket melon corn hamburger woman	restrictive
68	Si laki mankemema si bintang.	The man sews the BINTANG.	jacket melon apple fries man	restrictive
69	Si gadis mankemema si bintang.	The woman sews the BINTANG.	jacket melon cucumber ananas	restrictive
70	Si gadis mankemema si bintang.	The woman sews the BINTANG.	jacket bread cake chicken	restrictive
71	Si laki mankemema si bintang.	The man sews the BINTANG.	jacket mushroom pizza cheese man	restrictive
72	Si laki gumbumema si bintang.	The man points at the BINTANG.	skirt dress carrot sausage man	restrictive
73	Si gadis tambamema si awan.	The woman takes the AWAN.	jeans chicken corn hamburger woman	Non-restrictive
74	Si laki gumbumema si awan.	The man points at AWAN.	jeans chicken apple fries man	Non-restrictive
75	Si laki tambamema si awan.	The man takes the AWAN.	jeans chicken cucumber sausage	Non-restrictive
76	Si laki tambamema si awan.	The man takes the AWAN.	jeans bread banana ananas	Non-restrictive
77	Si gadis gumbumema si awan.	The woman points at the AWAN.	jeans mushroom cake cheese woman	Non-restrictive
78	Si gadis tambamema si awan.	The woman takes the AWAN.	dress skirt carrot melon woman	Non-restrictive
79	Si laki tambamema si ekor.	The man takes the EKOR.	shorts cheese corn hamburger man	Non-restrictive
80	Si laki gumbumema si ekor.	The man points at EKOR.	shorts cheese apple ananas man	Non-restrictive
81	Si laki tambamema si ekor.	The man takes the EKOR.	shorts cheese cucumber sausage	Non-restrictive
82	Si gadis gumbumema si ekor.	The woman points at the EKOR.	shorts bread banana melon	Non-restrictive
83	Si gadis tambamema si ekor.	The woman takes the EKOR.	shorts mushroom pizza chicken woman	Non-restrictive
84	Si gadis gumbumema si ekor.	The woman points at the EKOR.	coat scarf cake fries woman	Non-restrictive
85	Si gadis tambamema si lebah.	The woman takes the LEBAH.	cap ananas corn sausage woman	Non-restrictive
86	Si laki gumbumema si lebah.	The man points at LEBAH.	cap ananas carrot melon man	Non-restrictive
87	Si gadis tambamema si lebah.	The woman takes the LEBAH.	cap ananas cucumber chicken	Non-restrictive
88	Si laki gumbumema si lebah.	The man points at the LEBAH.	cap cake banana fries	Non-restrictive
89	Si laki tambamema si lebah.	The man takes the LEBAH.	cap hamburger apple cheese man	Non-restrictive
90	Si gadis gumbumema si lebah.	The woman takes the LEBAH.	apron hat pizza mushroom woman	Non-restrictive
91	Si gadis tambamemema si gula.	The woman takes the GULA.	gloves mushroom cucumber cheese woman	Non-restrictive
92	Si gadis gumbumema si gula.	The woman points at GULA.	gloves mushroom apple chicken woman	Non-restrictive
93	Si laki tambamema si gula.	The man takes the GULA.	gloves mushroom corn melon	Non-restrictive
94	Si laki gumbumema si gula.	The man points at the GULA.	gloves bread carrot fries	Non-restrictive
95	Si gadis tambamema si gula.	The woman takes the GULA.	gloves ananas banana sausage woman	Non-restrictive
96	Si gadis gumbumema si gula.	The woman points at the GULA.	vest socks pizza hamburger woman	Non-restrictive

Figure B.31.: Example picture noun testing, Exp. 5, Test Type 1
si sonis

Figure B.32.: Example picture noun testing, Exp. 5, Test Type 2
si sonis

Table B.11.: Vocabulary-Test Items Exp. 5

Item	Noun	Depicted Objects	Test Type	Condition
1	sonis	corn socks gloves scarf	1	Restrictive
2	sonis	hamburger socks jumper apron	2	Restrictive
3	bagus	apple hat shirt apron	1	Restrictive
4	bagus	fries hat top skirt	2	Restrictive
5	kemei	cucumber scarf jacket coat	1	Restrictive
6	kemei	sausage scarf jeans dress	2	Restrictive
7	towok	banana skirt shorts socks	1	Restrictive
8	towok	melon skirt cap coat	2	Restrictive
9	firel	pizza dress shirt hat	1	Non-restrictive
10	firel	chicken dress cap scarf	2	Non-restrictive
11	oblung	carrot coat jumper dress	1	Non-restrictive
12	oblung	cheese coat jacket socks	2	Non-restrictive
13	zelan	bread apron jeans vest	1	Non-restrictive
14	zelan	ananas apron shorts vest	2	Non-restrictive
15	pripas	cake vest gloves skirt	1	Non-restrictive
16	pripas	mushroom vest top hat	2	Non-restrictive
17	metin	shirt hamburger cake fries	1	Restrictive
18	metin	socks hamburger apple ananas	2	Restrictive
19	daram	jumper fries corn cheese	1	Restrictive
20	daram	hat fries cucumber chicken	2	Restrictive
21	kidir	top sausage banana melon	1	Restrictive
22	kidir	scarf sausage pizza cheese	2	Restrictive
23	bintang	jacket melon carrot chicken	1	Restrictive
24	bintang	skirt melon bread sausage	2	Restrictive
25	awan	jeans chicken corn mushroom	1	Non-restrictive
26	awan	dress chicken apple melon	2	Non-restrictive
27	ekor	shorts cheese cucumber ananas	1	Non-restrictive
28	ekor	coat cheese banana fries	2	Non-restrictive
29	lebah	cap ananas carrot hamburger	1	Non-restrictive
30	lebah	apron ananas cake mushroom	2	Non-restrictive
31	gula	gloves mushroom bread sausage	1	Non-restrictive
32	gula	vest musroom pizza hamburger	2	Non-restrictive

Bibliography

Aitchinson, J. (1987). *Words in the Mind: An Introduction to the Mental Lexicon.* Oxford and New York: Basil Blackwell, 2nd edition.

Akhtar, N. & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, 19(57), 347–358.

Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202–232.

Alishahi, A. & Fazly, A. (2010). Integrating syntactic knowledge into a model of cross-situational word learning. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual meeting of the Cognitive Science Society* (pp. 2452–2458). Austin, TX: Cognitive Science Society.

Allen, L. (1995). The Effects of emblematic gestures on the development and access of mental representations of French expressions. *The Modern Language Journal*, 79(4), 521–529.

Altmann, G. T. & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and visual world. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world.* New York: Psychology Press.

Altmann, G. T. M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–64.

Altmann, G. T. M. & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57(4), 502–518.

Amato, M. & MacDonald, M. (2010). Sentence processing in an artificial language: Learning and using combinatorial constraints. *Cognition*, 116(1), 143–148.

Arunachalam, S. & Waxman, S. R. (2010a). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, 114(3), 442–6.

Arunachalam, S. & Waxman, S. R. (2010b). Specifying the role of linguistic information in verb learning. In K. Franich, K. Iserman, & L. Keil (Eds.), *Proceedings of the 34th Annual Boston University Conference on Language Development.* (pp. 11–21). Somervile, MA: Cascadilla Press.

Au, T. K.-f. & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Trends in Cognitive Science*, 61(5), 1474–1490.

Baayen, R., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.

Baldwin, D. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 62, 875–890.

Baldwin, D. (1993). Infants' ability to consult the speaker for clues to word reference. *Journal of Child Language*, 20(2), 395–418.

Baldwin, D. (2000). Interpersonal understanding fuels knowledge aquisition. *Cartography and Geographic Information Science*, 9(2), 40–45.

Baldwin, D., 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), 3135–53.

Barcroft, J. (2004). Second language vocabulary acquisition: A lexical input processing approach. *Foreign Language Annals*, 37(2), 200–208.

Bates, D. (2005). Fitting linear mixed models in r. *R News*, (pp. 27–30).

Bates, E. & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218). New York: Cambridge University Press.

Bernal, S., Lidz, J., Millotte, S., & Christophe, A. (2007). Syntax constrains the acquisition of verb meaning. *Language Learning and Development*, 3, 325–341.

Bever, T. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–352). New York: Whiley.

Boland, J. E. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, & beyond* (pp. 51–76). Brighton, UK: Psychology Press.

Boland, J. E., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: Evidence from wh-questions. *Journal of Memory and Language*, 34, 774–806.

Brain, M. (1992). What sort of innate structure is needed to 'bootstrap' into syntax? *Cognition*, 45, 77–100.

Brandone, A. C., Pence, K. L., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Action speaks louder than words: young children differentially weight perceptual, social, and linguistic cues to learn verbs. *Child Development*, 78(4), 1322–42.

Breitenstein, C., Kamping, S., Jansen, A., Schomacher, M., & Knecht, S. (2004). Word learning can be achieved without feedback: implications for aphasia therapy. *Restorative Neurology and Neuroscience*, 22(6), 445–58.

Breitenstein, C. & Knecht, S. (2003). Language acquisition and statistical learning. *Der Nervenarzt*, 74(2), 133–43.

Breitenstein, C., Zwitserlood, P., de Vries, M. H., Feldhues, C., Knecht, S., & Dobel, C. (2007). Five days versus a lifetime: intense associative vocabulary training generates lexically integrated words. *Restorative Neurology and Neuroscience*, 25(5-6), 493–500.

Brown, G. (2008). Selective listening. *System*, 36(1), 10–21.

Bunger, A. (2006). How we learn to talk about events: Linguistic and conceptual constraints on verb learning. Unpublished PhD Thesis, Northwestern University.

Butterworth, G. (2004). Joint visual attention in infancy. In A. Slater & G. Bremner (Eds.), *Theories of infant development* (pp. 317–254). Oxford: Willey-Blackwell.

Carey, S. (1978). The child as a word learner. In J. Bresnan & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Carpenter, M., Nagell, K., Tomasello, M., Butterworth, G., & Moore, C. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, 63(4), (Serial No. 255).

Carter, R. (1987). *Vocabulary: Applied Linguistic Perspectives.* London: Allen and Unwin (2nd edition: Routledge, 1999).

Chambers, C. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language*, 47(1), 30–49.

Chambers, C. G., Tanenhaus, M. K., & Magnuson, J. S. (2004). Actions and affordances in syntactic ambiguity resolution. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 30(3), 687–96.

Chang, A. & Read, J. (2007). Support for foreign language listeners: Its effectiveness and limitations. *RELC Journal*, 38(3), 375–394.

Chiang, C. S. & Dunkel, P. (1992). The effect of speech modification, prior knowledge, and listening proficiency on EFL lecture learning. *TESOL Quarterly*, 26(2), 345.

Childers, J. B. & Paik, J. H. (2008). Korean- and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language*, 36(1), 201–24.

Christophe, A., Millotte, S., Bernal, S., & Lidz, J. (2008). Bootstrapping lexical and syntactic acquisition. *Language and Speech*, 51, 61–75.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. (2004). Phonological phrase boundaries constrain lexical access I. Adult data. *Journal of Memory and Language*, 51(4), 523–547.

Clahsen, H. & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, 27(01), 3–42.

Clark, J. M. & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210.

Cohen, A. (1987). The use of verbal and imagery mnemonics in second-language vocabulary learning. *Studies in Second Language Acquisition*, 9, 43–62.

Crocker, M. W., Knoeferle, P., & Mayberry, M. R. (2010). Situated sentence processing: the coordinated interplay account and a neurobehavioral model. *Brain and Language*, 112(3), 189–201.

Davis, M. H., Di Betta, A. M., Macdonald, M. J. E., & Gaskell, M. G. (2009). Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience*, 21(4), 803–20.

Dobel, C., Junghöfer, M., Breitenstein, C., Klauke, B., Knecht, S., Pantev, C., & Zwitserlood, P. (2010). New names for known things: on the association of novel word forms with existing semantic information. *Journal of Cognitive Neuroscience*, 22(6), 1251–61.

Dumay, N. & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological science: A Journal of the American Psychological Society*, 18, 35–39.

Elley, W. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning*, 41(3), 375–411.

Ellis, N. (1997). Vocabulary acquisition: Word structure, collocation, word-class, and meaning. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 122–139). Cambridge: Cambridge University Press.

Fahle, M. (2005). Perceptual learning: specificity versus generalization. *Current Opinion in Neurobiology*, 15, 1540–1602.

Federmeier, K. D. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology*, 44(4), 491–505.

Felser, C. & Clahsen, H. (2009). Grammatical processing of spoken language in child and adult language learners. *Journal of Psycholinguistic Research*, 38, 305–319.

Felser, C. & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research*, 23, 9–36.

Fender, M. (2001). A review of L1 and L2/ESL word integration skills and the nature of L2/ESL word integration development involved in lower-level text processing. *Language Learning*, 51, 319–296.

Fenn, K., Nusbaum, H., & Margoliash, D. (2003). Consolidation during sleep of perceptual learning of spoken language. *Nature*, 425, 614–616.

Ferreira, F. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25(3), 348–368.

Field, J. (2004). An insight into listeners' problems: too much bottom-up or too much top-down? *System*, 32(3), 363–377.

Field, J. (2007). Looking outwards, not inwards. *ELT Journal*, 61(1), 30–38.

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science*, 5(1), 55–64.

Fisher, C. & Gleitman, L. R. (2002). Language acquisition. In H. Pashler & C. Gallistel (Eds.), *Stevens' handbook of experimental psychology, Vol 1: Learning and motivation* (pp. 445–496). New York: Wiley.

Fisher, C., Klingler, S. L., & Song, H.-J. (2006). What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition*, 101(1), B19–29.

Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA: MIT Press.

Fodor, J. (1987). Modules, frames, fridgeons, sleeping dogs and the music of the spheres. In Z. Pylyshyn (Ed.), *The robot's dilemma: The frame problem in artificial intelligence* (pp. 139–149). Norwood, NJ: Ablex.

Frank, M., Goodman, N., & Tenenbaum, J. (2009). Using speaker's referential intentions to model early cross-situational word learning. *Psychological Science*, 20, 578–585.

Frazier, L. (1979). On comprehending sentences: Syntactic parsing strategies. Unpublished doctoral dissertation.

Frenck-Mestre, C. & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology*, 50A, 119–148.

Gangwani, T., Kachergis, G., & Yu, C. (2010). Simultaneous cross-situational learning of category and object names. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual meeting of the Cognitive Science Society* (pp. 1595–1601). Austin, TX: Cognitive Science Society.

Garnham, A. (1981). Mental models as representation of text. *Memory and Cognition*, 9, 560–565.

Gass, S. M. (1988). Second language vocabulary acquisition. *Annual Review of Applied Linguistics*, 9, 92–106.

Gernsbacher, M. A. & Kaschak, M. P. (2002). Language comprehension. In L. Nadel (Ed.), *Encyclopedia of cognitive sciences* (pp. 723–726). London, UK: Nature Publishing Group.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73(2), 135–76.

Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing*, 19(2), 133–167.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, 1(1), 3–55.

Gleitman, L. R. & Gleitman, H. (1992). A Picture Is worth at thousand words, but that's the problem: The role of syntax in vocabulary acquisition. *Current Directions in Psychological Science*, 1(1), 31–35.

Goh, C. (1997). Metacognitive awareness and second language listeners. *ELT Journal*, 51(4), 361–369.

Goldin-Meadow, S., Seligman, M., & Gelman, R. (1976). Language in the two-year old. *Cognition*, 4, 189–202.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, 28(1), 99–108.

Gollin, E. & Sharps, M. (1988). Facilitation of free recall by categorical blocking depends on stimulus type. *Memory and Cognition*, 16, 539–544.

Gómez, R., Bootzin, R., & Nadel, L. (2006). Naps promote abstraction in language-learning infants. *Psychological Science*, 17, 670–674.

Gout, A., Christophe, A., & Morgan, J. (2004). Phonological phrase boundaries constrain lexical access II. Infant data. *Journal of Memory and Language*, 51(4), 548–567.

Graham, S. A., Nilsen, E. S., Collins, S., & Olineck, K. (2010). The role of gaze direction and mutual exclusivity in guiding 24-month-olds' word mappings. *British Journal of Developmental Psychology*, 28(2), 449–465.

Gullberg, M., Roberts, L., & Dimroth, C. (in press). What word-level knowledge can adult learners acquire after minimal exposure to a new language? *International Review of Applied Linguistics*.

Gullberg, M., Roberts, L., Dimroth, C., & Veroude, K. (2010). Adult language learning after minimal exposure to an unknown natural language. *Language Learning*, 60(S2), 5–24.

Hamada, M. (2009). Development of L2 word-meaning inference while reading. *System*, 37(3), 447–460.

Hansen, M. & Markman, E. M. (2009). Children's use of mutual exclusivity to learn labels for parts of objects. *Developmental Psychology*, 45(2), 592–596.

Harrington, M. (2004). Between the input and the acquisition lies the shadow. *Bilingualism: Language and Cognition*, 7(1), 29–31.

Hayes, P. (1987). What the frame problem is and isn't. In Z. Pylyshyn (Ed.), *The robot's dilemma: The frame problem in artificial intelligence*. Norwood, NJ: Ablex.

Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40(1), 109.

Hoehle, B., Weissenborn, J., Kiefer, D., Schulz, A., & Schmitz, M. (2004). Functional elements in infants' speech processing: The role of determiners in the syntactic categorization of lexical elements. *Infancy*, 5, 341–353.

Hollich, G., Jusczyk, P. W., & Brent, M. R. (2000). How infants use the words they know to learn new words. In A. Do, L. Dominguéz, & A. Johansen (Eds.), *Proceedings of the 25th Annual Boston University Conference on Language Development.* (pp. 353–364). Somervile, MA: Cascadilla Press.

Horton, M. S. & Markman, E. M. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development*, 51(3), 708–719.

Houston-Price, C., Plunkett, K., & Duffy, H. (2006). The use of social and salience cues in early word learning. *Journal of Experimental Child Psychology*, 95(1), 27–55.

Howell, S., Jankowicz, D., & Becker, S. (2005). A model of grounded language acquisition: Sensorimotor features improve lexical and grammatical learning. *Journal of Memory and Language*, 53(2), 258–276.

Hulstijn, J. (1992). Retention of inferred and given word meanings: Experiments in incidental learning. In P. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113–125). London: Macmillan.

Hulstijn, J. (1997). Mnemonic methods in foreign language vocabulary learning: theoretical considerations and pedagogical implications. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 203–224). Cambridge: Cambridge University Press.

Hulstijn, J. (2001). Intentional and incidental second-language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258–286). Cambridge: Cambridge University Press.

Hulstijn, J. (2003). Connectionist models of language processing and the training of listening skills with the aid of multimedia software. *Computer Assisted Language Learning*, 16(5), 413–425.

Ichinco, D., Frank, M. C., & Saxe, R. (2009). Cross-situational word learning respects mutual exclusivity. In N. Taatgen, J. Van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual meeting of the Cognitive Science Society* (pp. 2214–2219). Austin, TX: Cognitive Science Society.

Johnson-Laird, P. (1983). *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness.* Cambridge: Cambridge University Press.

Jones, L. C. (2006). Effects of collaboration and multimedia annotations on vocabulary learning and listening comprehension. *CALICO Journal*, 24(1), 33–58.

Jones, L. C. & Plass, J. A. N. L. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal*, 86, 546–561.

Jones, S., Smith, L., & Landau, B. (1991). Object properties and knowledge in early lexical learning. *Child Development*, 62, 499–516.

Juffs, A. (2009). Second language acquisition of the lexicon. In W. Ritchie & T. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 181–211). Amsterdam, The Netherlands: Elsevier, 2nd edition.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2009a). Frequency and contextual diversity effects in cross-situational word learning. In N. Taatgen, J. Van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual meeting of the Cognitive Science Society* (pp. 2220–2225). Austin, TX: Cognitive Science Society.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2009b). Temporal contiguity in cross-situational statistical learning. In N. Taatgen, J. Van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual meeting of the Cognitive Science Society* (pp. 1704–1709). Austin, TX: Cognitive Science Society.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2010a). Adaptive constraints and inference in cross-situational word learning. In S. Ohlssen & R. Catrambone (Eds.), *Proceedings of the 32nd Annual meeting of the Cognitive Science Society* (pp. 1210–1215). Austin, TX: Cognitive Science Society.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2010b). Cross-situational statistical learning: Implicit or intentional? cross-situational statistical learning. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual meeting of the Cognitive Science Society* (pp. 1189–1194). Austin, TX: Cognitive Science Society.

Kaivanpanah, S. & Alavi, S. (2008). The role of linguistic knowledge in word-meaning inferencing. *System*, 36(2), 172–195.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49(1), 133–156.

Kelly, S., Barr, D. J., Breckenridge Church, R., & Lynch, K. (1999). Offering a Hand to pragmatic understanding: The role of speech and gesture in comprehension and memory. *Journal of Memory and Language*, 40(4), 577–592.

Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.

Kirkham, N., Slemmer, J., & Johnson, S. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition*, 83, B35–B42.

Klein, K. & Yu, C. (2009). Joint or conditional probability in statistical word learning: Why decide? In N. Taatgen, J. Van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual meeting of the Cognitive Science Society* (pp. 3105–3110). Austin, TX: Cognitive Science Society.

Knoeferle, P. & Crocker, M. (2007). The influence of recent scene events on spoken comprehension: Evidence from eye movements. *Journal of Memory and Language*, 57(4), 519–543.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: evidence from eye-movements in depicted events. *Cognition*, 95(1), 95–127.

Kotz, S. A. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and language*, 109(2-3), 68–74.

Krashen, S. (1985). *The Input Hypothesis*. London and New York: Longman.

Krashen, S. & Terrell, T. (1983). *The Natural Approach: Language Acquisition in the Classroom*. London: Prentice Hall Europe.

Kveraga, K., Ghuman, A. S., & Bar, M. (2007). Top-down predictions in the cognitive brain. *Brain and cognition*, 65(2), 145–68.

Lakoff, G. (1987). *Women, Fire, and Dangerous Things*. Chicago: The University of Chicago Press.

Landau, B. (1998). Object shape, object function, and object name. *Journal of Memory and Language*, 38(1), 1–27.

Landau, B. & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Landau, B., Smith, L., & Jones, S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, 3(3), 299–321.

Landau, B., Smith, L. B., & Jones, S. (1992). Syntactic context and the shape bias in children's lexical learning and adults. *Journal of Memory and Language*, 31, 807–825.

Langton, S., Law, A., Burton, A., & Schweinberger, S. (2008). Attentional capture of faces. *Cognition*, 107, 330–342.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In J. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126–132). London: Macmillan.

Laufer, B. & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: effects of language learning contex. *Language Learning*, 48, 365–391.

Lee, J. N. & Naigles, L. R. (2008). Mandarin learners use syntactic bootstrapping in verb acquisition. *Cognition*, 106(2), 1028–37.

Lew-Williams, C. & Fernald, A. (2007). How first and second language learners use predictive cues in online sentence interpretation in Spanish and English. In H. Caunt-Nulton, S. Kulatilake, & I. Woo (Eds.), *Proceedings of the 31st Annual Boston University Conference on Language Development* (pp. 382–393). Somervile, MA: Cascadilla Press.

Lidz, J., Bunger, A., Leddon, E., Baier, R., & Waxman, S. R. (2010). When one cue is better than two: Lexical vs. syntactic cues to verb learning. Unpublished manuscript.

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: verb learning and the footprint of universal grammar. *Cognition*, 87, 151–178.

Lindsay, S. & Gaskell, M. G. (2010). A complementary systems account of word learning in L1 and L2. *Language Learning*, 60, 45–63.

Llanes, A. & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, 37(3), 353–365.

Long, D. R. (1990). What you don't know can't help you. *Studies*, 12, 65–80.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution [corrected]. *Psychological Review*, 101(4), 676–703.

MacNamara (1972). Cognitive basis for language learning in infants. *Psychological Review*, 79, 1–13.

MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum Associates.

MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism* (pp. 49–67). Oxford: Oxford University Press.

MacWhinney, B. (2008). A unified model. In N. Ellis & P. Robinson (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 341–371). New York: Routledge.

Mandler, G. (1967). Organization and memory. In K. Spence & J. Spence (Eds.), *The psychology of learning and motivation: Advances in research and theory, vol. 1* (pp. 328–372). New York: Academic Press.

Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2003). Gaps in second language sentence processing. *Studies in Second Language Acquisition*, 27(01), 43–79.

Markman, E. (1990). Constraints children place on word meanings. *Cognitive Science*, 14(1), 57–77.

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

Markman, E. M. & Wachtel, G. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, 20, 121–157.

Markson, L., Diesendruck, G., & Bloom, P. (2008). The shape of thought. *Developmental science*, 11(2), 204–8.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244, 522–523.

Marslen-Wilson, W. & Tyler, L. (1987). Modularity in knowledge representation and natural language understanding. In J. Garfield (Ed.), *Modularity in knowledge representation and natural language understanding.* Cambridge, MA: MIT Press.

McNeill, D. (1992). *Hand and mind: What the hands reveal about thought.* Chicago, IL: University of Chicago Press.

McQuillan, J. & Krashen, S. (2008). Commentary: Can free reading take you all the way? A response to Cobb (2007). *About Language Learning & Technology*, 6, 104–109.

McRae, K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38(3), 283–312.

McRae, K. & Matsuki, K. (2009). People use their knowledge of common events to understand language, and do so as quickly as possible. *Language and Linguistics Compass*, 3, 1417–1429.

Medina, T., Hafri, A., Trueswell, J., & Gleitman, L. (2010). Propose but verify: Fast mapping meets cross-situational word learning. Paper presented at the Boston University Conference on Language Development (BUCLD) Annual Meeting, Boston, MA, Nov. 5.

Melka Teichroew, F. (1982). Receptive versus productive vocabulary: a survey. *Interlanguage Studies Bulletin*, 6(2), 5–33.

Mendelsohn, D. (2001). Listening comprehension: We've come a long way, but... *Contact*, 27, 33–40.

Meskill, C. (1996). Listening skills development through multimedia. *Journal of Multimedia and Hypermedia*, 5(2), 179–201.

Meyer, A., Sleiderink, W., & Levelt, W. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, 66, B25–B33.

Millotte, S., Wales, R., Dupoux, E., & Christophe, A. (2006). Can prosodic cues and function words guide syntactic processing and acquisition? In R. Hoffmann & H. Mixdorff (Eds.), *Speech prosody: 3rd international conference*. Dresden: Dresden TUD Press.

Milton, J. & Meara, P. (1995). How periods abroad affect vocabulary growth in a foreign language. *International Review of Applied Linguistics*, 108, 17–34.

Monaghan, P. & Mattock, K. (2009). Cross-situational language learning: The effects of grammatical categories as constraints on referential labeling. In N. Taatgen, J. Van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual meeting of the Cognitive Science Society* (pp. 2226–2231). Austin, TX: Cognitive Science Society.

Mondria, J.-A. & Wit-De Boer, M. (1991). The effects of contextual richness on the guessability and the retention of words in a foreign language. *Applied Linguistics*, 12(3), 249–267.

Moore, C., Angelopoulos, M., & Bennett, P. (1999). Word learning in the context of referential and salience cues. *Developmental psychology*, 35(1), 60–8.

Morgan, J. & Demuth, K. (1996). Signal to syntax: An overview. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 1–22). Mahwah, NJ: Lawrence Erlbaum Associates.

Mueller, J. (1980). Visual contextual cues and listening comprehension: An experiment. *Modern Language Journal*, 64(3), 335–340.

Mueller, J. (2005). Electrophysiological correlates of second language processing. *Second Language Research*, 21, 152–174.

Mueller, J. L. (2009). The influence of lexical familiarity on ERP responses during sentence comprehension in language learners. *Second Language Research*, 25(1), 43–76.

Naigles, L. & Swensen, L. (2007). Syntactic supports for word learning. In E. Hoff & M. Shatz (Eds.), *The handbook of language development* (pp. 212–231). New York: Blackwell.

Nappa, R., Wessel, A., McEldoon, K., Gleitman, L., & Trueswell, J. (2009). Use of speaker's gaze and syntax in verb learning. *Language Learning and Development*, 5(4), 203–234.

Nation, I. (1982). Beginning to learn foreign vocabulary: A review of the research. *RELC Journal*, 13(1), 14–36.

Nation, I. (1990). *Teaching and Learning Vocabulary.* New York: Newbury House.

Nation, I. (2001). *Learning Vocabulary in Another Language.* Cambridge: Cambridge University Press.

Nation, I. & Waring, R. (1978). Vocabulary size, text coverage, and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 6–19). New York: Cambridge University Press.

Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP study of second language learning after childhood: effects of proficiency. *Journal of Cognitive Neuroscience*, 17, 1212–1228.

Papadopoulou, D. & Clahsen, H. (2003). Parsing in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, 25, 501–528.

Paribakht, T. S. & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary development. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174–200). New York: Cambridge University Press.

Paribakht, T. S. & Wesche, M. (1999). Reading and incidental L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition*, 21(2), 195–224.

Paribakht, T. S. & Wesche, M. (2010). *Lexical Inferencing in a First and Second Language: Cross-linguistic Dimensions.* Clevedon, Avon, UK: Multilingual Matters.

Pickering, M. J., Clifton, C., & Crocker, M. W. (1999). Architectures and mechanisms in sentence comprehension. In M. Crocker, M. J. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 1–31). Cambridge: Cambridge University Press.

Pickering, M. J. & Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3), 105–10.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second language multimedia learning environment. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 90(1), 25–36.

Prince, P. (1996). Language vocabulary the role of context learning: Function of proficiency. *The Modern Language Journal*, 80(4), 478–493.

Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Language Learning*, 53(2), 233–284.

Quine, W. (1960). *Word and Object*. Cambridge, MA: Cambridge University Press.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. *Journal of Verbal Learning and Behavior*, 22, 358–374.

Read, J. (2000). *Assessing Vicabulary*. Cambridge: Cambridge University Press.

Read, J. (2004). Research in teaching vocabulary. *Annual Review of Applied Linguistics*, 24, 146–161.

Richards, J. (2000). Series editor's preface. In N. Schmitt (Ed.), *Vocabulary in language teaching* (pp. xi–xii). Cambridge: Cambridge University Press.

Roberts, L. & Felser, C. (2011). Plausibility and recovery from garden-paths in second-language sentence processing. *Applied Psycholinguistics*, 32, 299–331.

Rost, M. (2002). *Teaching and Researching Listening*. London: Longman.

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78(2), 199–221.

Sabourin, L. & Haverkort, M. (2003). Neural substrates of representation and processing of a second language. In R. Van Hout, A. Hulk, F. Kuiken, & R. Towell (Eds.), *The lexicon-syntax interface in second language acquisition* (pp. 175–195). Amsterdam: John Benjamins.

Sagarra, N. & Alba, M. (2006). The Key is the Keyword: L2 Vocabulary Learning Methods with Beginning Learners of Spanish. *The Modern Language Journal*, 90, 228–243.

Sagiv, N. & Bentin, S. (2001). Structural encoding of human and schematic faces: Holistic and part based processes. *Journal of Cognitive Neuroscience*, 13, 1–15.

Schmitt, N. (1997). Vocabulary learning strategies. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 199–227). New York: Cambridge University Press.

Schmitt, N. (2000). *Vocabulary in Language Teaching*. New York: Cambridge University Press.

Schoonen, R., Hulstijn, J., & Bossers, B. (1998). Metacognitive and language-specific knowledge in native and foreign language reading comprehension: An empirical study among Dutch students in grades 6, 8 and 10. *Language Learning*, 48(1), 71–106.

Schwanenflugel, P. & Shoben, E. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language*, 24, 232–252.

Secules, T., Herron, C., & Tomasello, M. (1992). The effect of video context on foreign language learning. *The Modern Language Journal*, 76(4), 480–490.

Sedivy, J. C., Tanenhaus, M. K., Chambers, C. G., & Carlson, G. N. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, 71(2), 109–47.

Seo, K. (2002). The effect of visuals on listening comprehension: A study of Japanese learners' listening strategies. *Internation Journal of Listening*, 16, 57–82.

Sharwood Smith, M. (1991). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15, 165–179.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA. *Studies in Second Language Acquisition*, 15, 165–179.

Shatzman, K. B. & McQueen, J. M. (2006). Prosodic knowledge affects the recognition of newly acquired words. *Psychological Science*, 17(5), 372–7.

Shepherd, S. (2010). Following gaze: gaze-following behavior as a window into social cognition. *Frontiers in Integrative Neuroscience*, 4, 5.

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61(1-2), 39–91.

Smith, K., Smith, A. D. M., & Blythe, R. A. (2010). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science*, (pp. 1–19).

Smith, L. & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, 106(3), 1558–68.

Snedeker, J. (2009). Word learning. In L. Squire (Ed.), *The new encyclopedia of neuroscience* (pp. 503–508). Amsterdam: Elsevier.

Snedeker, J. & Gleitman, L. (2004). Why is it hard to label our concepts. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 257–295). Cambridge: MIT Press.

Snow, C. (1999). Social perspectives on the emergence of language. In B. MacWhinney (Ed.), *The emergence of language* (pp. 257–276). Mahwah, NJ: Lawrence Erlbaum Associates.

Spivey, M. J., Tanenhaus, M. K., Eberhard, K. M., & Sedivy, J. C. (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, 45(4), 447–81.

St. Clair, M. C. & Monaghan, P. (2008). Language abstraction: Consolidation of language structure during sleep. In B. Love, K. McRae, & V. Sloutsky (Eds.), *Proceedings of the 30th Annual meeting of the Cognitive Science Society* (pp. 727–733). Austin, TX: Cognitive Science Society.

Staub, A. & Clifton, C. (2006). Syntactic prediction in language comprehension: evidence from either...or. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 32(2), 425–36.

Sueyoshi, A. & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, 55(4), 661–699.

Tamminen, J. & Gaskell, M. G. (2008). Newly learned spoken words show long-term lexical competition effects. *Quarterly Journal of Experimental Psychology*, 61, 361–371.

Tanenhaus, M. K. & Trueswell, J. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Handbook of perception and cognition, vol 11: Speech, language, and communication* (pp. 217–262). San Diego, CA: Academic Press, 2nd edition.

Tight, D. G. (2010). Perceptual learning style matching and L2 vocabulary acquisition. *Language Learning*, 60(4), 792–833.

Tomalski, P., Csibra, G., & Johnson, M. (2009). Rapid orienting toward face-like stimuli with gaze-relevant contrast information. *Perception*, 38(4), 569–578.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, 74(3), 209–53.

Tomasello, M. & Barton, M. (1994). Learning words in nonostensive contexts. *Developmental Psychology*, 30(5), 639–650.

Tomasello, M. & Todd, J. (1983). Joint attention and lexical acquisition style. *First Language*, 4, 197–212.

Trueswell, J., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285–318.

Tulving, E. & Gold, C. (1963). Stimulus information and contextual information as determinants of tachistoscopic recognition of words. *Journal of Experimental Psychology*, 66, 319–327.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language*, 41(2), 147–182.

Van Berkum, J. J. a., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(3), 443–467.

Van Gompel, R. & Pickering, M. J. (2007). Syntactic parsing. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 455–503). Oxford: Oxford University Press.

Vandergrift, L. (2004). Listening to learn or learning to listen? *Annual Review of Applied Linguistics*, 24, 3–25.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(03), 191.

Vogt, P. & Smith, A. (2005). Learning colour words is slow: A cross-situational learning account. *Behavioral and Brain Sciences*, 28, 509–510.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, 107(2), 729–42.

Vouloumanos, A. & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental psychology*, 45(6), 1611–7.

Wagner, U., Gais, S., Haider, H., Verleger, R., & Born, J. (2004). Sleep inspires insight. *Nature*, 427, 352–355.

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30(01), 79–95.

Wesche, M. & Paribakht, T. S. (2010). *Lexical Inferencing in a First and Second Language: Cross-linguistic Dimensions*. Clevedon, Avon, UK: Multilingual Matters.

Williams, J. N. (2006). Incremental interpretation in second language sentence processing. *English*, 9(1), 71–88.

Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2007). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive psychology*, 56(3), 165–209.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text. *Language Learning & Technology*, 13(2), 48–67.

Yeh, Y. & Wang, C.-w. (2003). Effects of multimedia vocabulary annotations and learning styles on vocabulary learning. *Trends in Cognitive Science*, 21(1), 131–144.

Yu, C. & Ballard, D. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, 70, 2149–2165.

Yu, C. & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological science : A journal of the American Psychological Society*, 18(5), 414–20.

Yuan, S. & Fisher, C. (2009). 'Really? She blicked the baby?' - Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, 20(5), 619–627.

Yurovsky, D., Fricker, D., Yu, C., & Smith, L. B. (2010). The active role of partial knowledge in cross-situational word learning. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual meeting of the Cognitive Science Society* (pp. 2609–2615). Austin, TX: Cognitive Science Society.

Yurovsky, D. & Yu, C. (2008). Mutual exclusivity in cross-situational statistical learning. In B. Love, K. McRae, & V. Sloutsky (Eds.), *Proceedings of the 30th Annual meeting of the Cognitive Science Society* (pp. 715–720). Austin, TX: Cognitive Science Society.

Zwitserlood, P., Klein, W., Liang, J., Perdue, C., Kellerman, E., & Wenk, B. (1994). The first minutes of foreign-language exposure. Unpublished manuscript.