The Effect of Different Teaching Techniques on Acquiring the Grammatical Gender of Nouns in German as a Foreign Language

---

Jessica Arzt

University of Alberta

Honors Thesis 2013

Table of Contents

Introduction:

The grammatical gender of German nouns poses one of the most difficult and persistent challenges to second language learners. Not only must one learn and identify the gender of each noun, but a language learner must also learn how each of the three genders, masculine, feminine, or neuter, affects other language elements so everything coordinates in a grammatically correct utterance. Even in advanced learners, gender mistakes make up a large portion of errors (Rogers, 1987, p. 49) and beginning learners may be entirely confused by the seemingly arbitrary assignment of gender. Especially for learners with an L1 like English that does not assign a grammatical gender, learning and applying this new grammatical concept can prove to be particularly difficult. However, knowing the grammatical gender of a noun (referred to simply as gender from here on) is essential in formulating correct sentences, as gender affects various morphosyntactic elements and cannot be separated from case and number (Rogers, 1987, p. 49). Gender is closely bound to syntactic considerations of even the most basic sentences, and learning and understanding this grammatical element is an essential step in becoming proficient in German.

While gender may be considered one of the harder concepts in German, it is one all learners are faced with right from the beginning. The first nouns a learner encounters, and every noun after that, come linked with a grammatical gender. Unlike other grammatical concepts though that can be learned based on a few consistent rules, gender is unique and difficult in its lack of transparency, and rules often look more like general patterns or complex algorithms. To a beginning learner, as well as for many advanced and native speakers, gender assignment seems entirely arbitrary and confusing. However, as Salmons outlines, the theoretical principles of parsimony and economy dictate that it is unlikely that a language would assign each and every

noun a gender without any rhyme or reason (Salmons, 1993, p. 417). This has led to much debate in literature about whether gender information is stored in the mental lexicon as one unit along with the representation of a noun, or whether a rule system for gender assignment exists in the mind that allows gender to be combined with a word upon retrieval (Delisle, 1985; Levine, 1999; Menzel & Tamaoka, 1995; Rogers, 1987; Salamoura, 2007; Schmidt, 1990). These opposing views suggest on the one hand that gender must be memorized along with lexical items and on the other hand that gender can be acquired by learning a system of rules. This debate has not only influenced theoretical research but has also been important to questions of applied linguistics and second language acquisition and has been influential to the question "How does one best teach German gender?"

Literature Review:

Important background considerations for this study include firstly a look at psycholinguistic models and approaches to the mental organization of gender as a part of language acquisition, followed by an examination of theories of memory and learning as they apply to language, and lastly a review of previous research on techniques for teaching grammatical gender.

Mental organization of gender:

An understanding of how language is stored and processed from a psycholinguistic perspective is fundamental to formulating effective teaching techniques. A direct access vs. a rule-based analysis of gender assignment suggests different mental organizations of language and in turn different approaches to teaching gender. Arguments in favor of rules have been

strong. Theoretical arguments like Salmons' posit that "a rule system, even one with relatively many exceptions, is far preferable to marking each noun and affix in the lexicon individually for gender" (Salmons, 1993, p. 417). This argument strongly appeals to logic. Other examinations argue that native speakers seem to have a "feel" for gender, and even language learners employ strategies when assigning gender (Delisle, 1985; Levine, 1999; Menzel & Tamaoka, 1995). In studies in which native speakers are asked to assign a gender to borrowed words or nonsense nouns, there is clear evidence of above chance agreement on gender assignment (Delisle, 1985; Levine, 1999). These results have been produced consistently and convincingly in a number of studies. This seems to suggest that gender assignment is not completely arbitrary and that some set of patterns, rules, or criteria exist in the minds of native speakers, which determine a particular gender. If this were not true, a German speaker would not be able to identify the gender of a noun he or she has not previously come across, which is clearly untrue. However, the majority of native speakers "cannot make explicit the rules" they are applying, indicating that this knowledge is not consciously employed when assigning a gender (Rogers, 1987, p. 55). Furthermore, language learners with lower language proficiency do not seem to show these consistencies in gender assignment (Menzel & Tamaoka, 1995), and Rogers (1987) suggests that even learners who can articulate specific gender rules cannot always apply them consistently. This seems to problematize the importance of rules. While Delisle (1985) found that even language learners with various proficiency levels did not assign gender completely arbitrarily to unknown loan words, their strategies did not mirror those of native speakers. Instead, strategies seemed to attempt to link German gender assignment to gender classes used in English, or culturally in general, and ultimately it was concluded that beginning learners view gender as random. Menzel and Tamaoka (1995) also suggest that even in children learning German as

their native language, these rules are not factored in at first. Rather, children first learn a noun and its gender as an amalgam and then slowly begin to pick up rules beginning with the most consistent rules and continuing to store certain exceptions as units. Rogers summarizes that "explicit knowledge of grammatical rules is therefore neither necessary (consider the native speaker) nor sufficient (consider the L2 learner) for correct usage" (Rogers, 1987, p. 55). This research seems to indicate that at some level of language processing rule-like organization contributes to gender assignment, but the value of these rules to learning may be questionable.

Various attempts have been made to articulate a set of gender assignment rules that try to make the system more transparent and to define the criteria by which Germans assign gender. Consistently, these rule frameworks seem to indicate patterns of gender assignment based on semantic, phonological, and morphological features (Delisle, 1985; Köpcke & Zubin, 1984; Salmons, 1993). These rules also function hierarchically with semantic features tending to come out strongest. However, rules are often not definite and lead to an either-or option or a further rule, and even at high levels of complexity, a completely comprehensive formula for gender assignment does not seem to exist. These rule systems also all seem inadequate with learners in mind. Those approaches that succeed at covering a greater percentage of nouns end up with an extensive list of complicated rules, and those approaches that attempt to compile a concise and simple list only manage to cover an inadequate number of words (Rogers, 1987). Bearing in mind that there are a few clear and consistent rules that are helpful to learners, on the whole, Rogers (1987) emphasizes that these rule approaches have little pedagogical value, with rules being complicated, unclear, and too numerous to be applied in any communicative setting by a learner.

If a rule-like system seems to underlie gender assignment by native speakers, but describing those rules is not particularly valuable for second language acquisition, the possibility of a more useful framework may lie in other theoretical approaches. Levine (1999) and Salmons (1993) suggest the value of a dual-route model for explaining gender assignment. This approach argues that in native speakers, frequent words in a language "are stored independently with all of their morphological [and gender] baggage," while other words must first be processed and assembled in order to be accessed (Salmons, 1993, p. 428). This means that frequent German nouns would be stored with their gender information as one unit, and less frequent words are combined with their gender based on rules while being retrieved. This is valuable, as frequent words in a language often break the rules. Menzel and Tamaoka (1995) suggest that in children learning German as an L1, or any gender assigning language, all nouns are learned as amalgams at first with gender and word as one unit. As vocabulary grows and frequency patterns and rules begin to emerge, the mental lexicon is reorganized and access to particular words is distinguished. This may be comparable to the way second language learners acquire gender. In this way, this model stands in support of both a direct-access approach as well as a rule-based system with language proficiency and word frequency as important variables. Such a model has various implications, but important to this study, seems to suggest that early and important words are stored and accessed as units with their gender information rather than by rules, and that beginning learners approach noun and gender as an amalgam.

An approach that comes at the problem from a different angle is a usage-based or connectionist model. Here the mind is compared to a computer-like information processor. Rather than explicit knowledge, language information is represented by a network of nodes linked by differently weighted connections that lead to a specific output in response to a

particular input (Hoff, 2009). The approach posits that language does not function over abstract symbols and rules at all. Language input is fundamental, and a language processor is built up as patterns are found and connections are formed that act like rules to produce a specific result in response to a certain kind of input, but no abstract rules or constraints actually exist (Hoff, 2009). Importantly, it is argued that, "connectionism is a mechanism for extracting regularities from experience, that is, for acquiring knowledge that is not directly given in any single experience" (Hoff, 2009, p. 21). Instead of explicit learning, acquisition occurs via experience and input as a whole. Schmidt furthers that learning involves "the gradual accumulation of associations between frequently co-occurring features [such as noun and gender], rather than unconscious induction of abstract rule systems" (Schmidt, 1990, p. 149). Each language experience contributes to the development of patterns of association that are systematically applied depending on stronger and weaker connections, and new language situations are approached based on experience with past situations that fit a pattern.

Connectionism has found much support in research on first language acquisition and seems to have value for second language acquisition as well. This model shows how gender assignment can be considered not arbitrary, as has been shown by native speakers, but that it also does not follow a set of rules that can be easily articulated. Instead, criteria are derived from experiences and are applied to produce output consistently as if by rules, without explicit knowledge of actual rules. Such a model accounts for initial mistakes and inconsistencies learners make, but there is an eventual build up to a reliable gender assignment system through exposure to more input. Furthermore, it is a framework that explains a way in which unfamiliar words can be approached that consistently produces the same output among different speakers. If one assumes that native speakers of a language have built up a relatively similar network of

mental connections, a fairly uniform structure exists by which a specific gender might be consistently assigned to an unfamiliar word with some room for variation. An experiment such as Levine's (1999) in which nonsense nouns are assigned a gender could be supported through such a model, showing indicative trends of gender assignment but with some native speakers displaying different assignment patterns. Such a model also accounts for the notion suggested by researchers such as Menzel and Tamaoka (1995) and Salmons (1993) that within the rule system there is a certain hierarchy and interdependence between rules that gives some rules more weight and influence over others. In a connectionist model, the network functions on weighted connections, rather than rules, which produce a specific output based on various interconnected aspects. If a particular association appears or applies more often in input, it will more strongly suggest a particular outcome. Semantic considerations of a word may create stronger associations and have a larger effect on gender assignment. In a connectionist model, mental rules are replaced by patterns of association gradually built up in a language processor. A speaker is not aware of the functioning of this processor or the algorithm it abides by, only the output it produces. In such a way, a connectionist approach models how native language processing happens, how learning builds up to such a framework, and why rule-like language behavior can be observed.

A connectionist model also has significant value in a pedagogical context. It suggests that a certain basis of learned vocabulary must first exist from which patterns can then be subconsciously determined (Menzel & Tamaoka, 1995, p. 20). As a learner is exposed to more input, pattern hypotheses become more accurate and gender assignment becomes more proficient and native-like. The distinction between direct and rule-based access is blurred here, and input and patterns are of central importance. This means that explicit instruction on gender assignment

rules may benefit learning if they are easy and consistent (for example learning that words ending in *e* are most often feminine) and may help strengthen certain mental associations, but they are by no means necessary for learning. Instead, exposure to input is essential and learners slowly get better at gender assignment as they acquire more of the language. From a pedagogical perspective, such a model of language learning means that especially in beginning learners of German, new lexical entries are stored with noun and gender as one unit and need to be learned and memorized. Only after a basis of words has been acquired, can patterns slowly be built up and subconsciously applied.

Schmidt summarizes, that "if an emerging consensus can be identified, it is in support of connectionist and other memory-based models that emphasize the importance of specific knowledge of instances over abstraction" (Schmidt, 1993, p. 215). With such a psycholinguistic model as a basis for an applied approach to second language learning, consideration must be turned to questions of how specific instances of words and gender are taken up by the learner.

Memory and Learning:

The psychology of memory and learning provides further fundamental considerations for effective teaching of gender. It has been argued above that input is critical for the acquisition of gender by a language learner. However, it is important that a distinction be made between input and intake. Sharwood Smith defines input as "language data that the learner is exposed to, that is...the potentially processible language data which are made available, by chance, or by design, to the language learner" (Sharwood Smith, 1993, p. 167). Intake on the other hand is "that part of input that has actually been processed by the learner and turned into knowledge of some kind" (Sharwood Smith, 1993, p. 167). In this sense, exposure to language input in the broadest sense

is inadequate; just because environmental input exists, does not mean the learner is taking it up. In a language learner, "language proficiency either develops as a response to input or fails to grow despite that input" (Sharwood Smith, 1993, p. 167). What makes the difference is whether the learner pays attention to the input being presented to him or her. Psychological models of memory as well as theories of second language acquisition have argued that attention is the key to turning input into intake and processing the information further to move it into long term memory storage (Schmidt, 1993). Gass (1997) outlines an input-interaction perspective for second language acquisition, which concentrates on the negotiation of meaning as a mode by which learners interact with input and focus attention on unfamiliar aspects in order to integrate them into a developing interlinguistic system (p. 87). An information processing perspective also emphasizes the importance of the attentional demands of filtering input and that learning occurs as attention is drawn to gaps between the learner's interlanguage system and the target system (Gass, 1997, p. 92). Gass and Selinker (2001) further outline the crucial role of attention to form in moving information from language input to intake and finally to output and the positive effects evident if attention to form occurs as information is initially processed (p. 315). These models all regard attention as an essential component in the acquisition of formal aspects of language.

Awareness does not always occur to the same degree, and different levels of attention result in a different interaction with input. Schmidt (1990) distinguished between three levels of awareness. Perception is the most surface-level form of attention; noticing is "the basic sense in which we commonly say that we are aware of something" (Schmidt, 1990, p. 132) and the "necessary and sufficient condition for the conversion of input to intake" (Schmidt, 1993, p. 209); understanding is the deepest level of awareness at which reflection and analysis of the

objects of consciousness occurs (Schmidt, 1990, p. 132). For language learning, noticing and understanding are of greatest significance and allow processing at a deep enough level to contribute to the acquisition of language. Schmidt elaborates that, "noticing is related to rehearsal within working memory and the transfer of information to long-term memory, to intake, and to item learning. Understanding is related to the organization of material in long term memory, to restructuring, and to system learning" (Schmidt, 1993, p. 213). In relation to the learning of gender, this means that the gender of a noun must not just exist in input presented to a learner. Actually processing and storing gender requires the learner to pay enough attention to notice the information in the input, and beginning to learn regularities and patterns requires deeper analysis and awareness at the level of understanding.

Information processing theories emphasize that awareness must be viewed as a limited mental capacity. This can be confirmed by anyone who has attempted to multi-task without success. Attention can be considered a type of "switchboard, gate, or filter that prevents us from being overwhelmed by the complexity of input" (Schmidt, 1990, p. 136). Select aspects of the total available input are allotted more or less of a limited resource of attention. More attention or awareness leads to deeper processing and a transfer of information first to short term memory and then with even further processing to long term memory (Schmidt, 1990). Those aspects of input not attended to, go unnoticed, and do not become intake. Attention and awareness are therefore essential to memory and in turn to language learning and specifically gender learning. The amount of attention a process requires is dependent on its level of automaticity. New skills demand a high level of attention and "cannot be carried out concurrently with other demanding tasks" (Schmidt, 1990, p. 136). As tasks or skills become easier through practice, they can become more automatic and occur parallel to other tasks. This suggests that a beginning

language learner is faced with a multitude of new and demanding tasks that cannot all be attended to equally. A certain competition for attention requires the information processor to make selections as to what is deemed most important until skills become automatic. A skill such as gender assignment and usage may be a relatively demanding skill for a beginning learner and may not always receive the required attention needed for acquisition.

It is important to distinguish that the teacher may be in control of input, but only the learner has control over attention, noticing, and intake. Sharwood Smith (1993) has proposed the notion of input enhancement in instructed second language acquisition as a method by which the saliency of certain aspects of input can be increased by the teacher to increase the likelihood it becomes intake for the learner. It is argued that input can be manipulated so that learner attention can possibly be steered in a specific direction and learner knowledge can be optimally affected. Such manipulation attempts to flag aspects of input deemed important by the teacher so the student may focus their attention, create internal mental "flags", and advance learning (Sharwood Smith, 1991, p. 120). Input enhancement can take various forms from unelaborated manipulation of font to explicit focus on grammar. Rather than leaving it entirely up to the learner to focus attention, the teacher attempts to emphasize what is important and capture the learner's attention.

As described by Sharwood Smith, input enhancement focuses specifically on aspects of form and has been applied as an aid to grammar learning in numerous experiments. Attention has further been argued to be essential to noticing formal elements of language and making adjustments to a developing interlanguage system. While gender is in fact an element of German grammar, it is, as mentioned previously, difficult in that it is closely bound to words and the mental lexicon. It was argued that gender is not something that is initially learned systematically

but rather is acquired by learning specific instances of input and building up a system gradually. If a theoretical framework is assumed in which beginning learners must memorize a noun and its respective gender as a lexical unit, the question stands if and how input enhancement can also be effective for the acquisition of gender and aid students in noticing this formal information in language input and learning it along with lexical information.

Teaching Techniques:

Neither simple exposure to input nor rote memorization of vocabulary lists are the only answer to the acquisition of gender. Just as with the learning of vocabulary, various teaching techniques have been suggested to make learning gender easier for students. However, unlike for vocabulary teaching and learning techniques, one must keep in mind that gender is a grammatical concept. Likely, at least initially, the information is stored in the mental lexicon and not in the mental grammar, but unlike its lexical counterpart, gender has no significant meaning affecting the direct understanding of a word – which for many learners may initially be the greatest area of focus. Furthermore, in actual communication, the gender of a noun never appears as *feminine, masculine,* or *neuter.* Rather, gender takes on a series of forms as either *die, der,* or *das,* or any forms of the definite or indefinite article or on other morphological markings. How then can gender best be represented to make learning easiest? Further, in respect to the previous consideration of input enhancement, numerous possibilities exist as manipulations by which gender can possibly be made more salient. Which of these possibilities has the greatest impact and is most effective for learners of various ages and proficiency levels is unclear.

The mnemonic keyword method has been suggested as an effective technique to help students learn foreign vocabulary. As suggested by Atkinson (1975), this technique involves

choosing a keyword in the learner's native language which is orthographically similar to the foreign word to be learned. An interactive image linking the keyword and the target word referents is then thought up by the learner. For example, to learn the German word *der Hund*, a learner may use the orthographically similar *hand* as a keyword and imagine a hand petting a dog. Such "acoustic" and "imagery" links are meant to help associate form and meaning and thereby support later recall by the learner (Atkinson, 1975, p. 821). In a study by Desrochers, Gelinas, and Wieland (1989) and a second study by Desrochers, Wieland, and Cote (1991) the mnemonic keyword method was adapted to help facilitate the learning of grammatical gender along with word meaning. Their modifications to the keyword method consisted of adding a third factor in the form of a gendered actor correlating to the gender of the noun to be learned, i.e. a man, a woman, or an inanimate object for words that were respectively masculine, feminine, or neuter. In this way, grammatical gender was linked to familiar concepts of sex. In experimental conditions, students were asked to think up an image in which the target word referent, the keyword referent, and a corresponding actor representing the grammatical gender were all somehow interacting (Desrochers et al., 1989). The idea behind the modification was to "[recode] the gender tags to make them more concrete and thereby enable students to form a semantic link between the referent of the German noun and the recoded gender tag" (Desrochers et. al., 1989, p. 25). Grammatical gender was turned into something imaginable and tangible, and word meaning, form, and gender all become linked in an associative image that aimed to make recall of all elements easier.

Both the studies by Desrochers et al. (1989) and Desrochers et al. (1991) came to fairly decisive conclusions. It was found that the modified keyword method did in fact facilitate the recall of gender. This was attributed to the "inclusion of a concrete representation of the recoded

gender tag in the interactive mental image" (Desrochers et al., 1991, p. 21). However, a key factor in recall was that the translation of the target word also be recalled. When students could not remember the meaning of the target word, gender assignment was only at chance levels. This seems to indicate a mental unit of noun meaning and gender that are stored and recalled together. Morphological markings of the target word alone were not enough to facilitate gender recall. The word meaning and semantic considerations were essential and recall of translation and gender were closely linked. Furthermore, it was tested if the effectiveness of the mnemonic keyword method changed when only the two elements linking lexical information were associated in comparison to when a third gender element was added. Here Desrochers et al. (1989) found that the additional load of learning gender resulted in a significant decrease in learning the translation of the target word. Possibly the task of coming up with such a complex mental image associating so many elements was too great a load on the learner. Interestingly, a further experiment was conducted in which students were instructed on the order in which to learn the gender and translation. It was found that the recall of gender was significantly better when students were instructed to learn the gender of the target noun before focusing on its meaning (Desrochers et al., 1991, p. 33). This seems to suggest that not only is there a significant load on memory and information processing by using this complex mnemonic strategy, but the unfamiliar concept of grammatical gender adds a further load that, in combination, may be exceedingly demanding for the learner. In respect to models that depict attention as a limited resource, the demands of this modified keyword technique must be critically considered.

Desrochers et al. (1991) also conducted one experiment in which students were instructed to apply the modified keyword method only to the noun gender, only to the translation, or to

apply the keyword method to both gender and translation. In this experiment, it was found that the keyword method facilitated the recall of gender whether or not word translation was also part of the interactive imagery. This suggests the possibility of the effectiveness of applying the method only to gender and thereby potentially reducing the load of the technique on mental processing while still attaining the benefits of the technique for gender acquisition.

This technique seems valuable for several reasons. It focuses on learning the noun and its gender as one unit without looking to rules for assignment. It enhances the gender information in the input making it salient to the student and encouraging its uptake and association in the mental lexicon. It gives a basis for association that links the possibly unclear and abstract category of noun gender to a more concrete concept of gender or sex, and aids the organization of information in memory to make later recall easier.

A different approach to enhancing gender information in language input was taken in studies by Nyikos (1987) and Kohler (2009). These studies used color-coding as a teaching technique to increase the saliency of grammatical gender and elaborate it in such a way to help categorize the information upon intake and facilitate recall. In the studies, each of the three genders was assigned a color and language input was coded accordingly. While Kohler (2009) focused on color coding nouns and gender information in a variety of grammatical cases, both in context and stand-alone vocabulary, Nyikos (1987) color-coded vocabulary words and pictures either alone or in combination. Both studies found a significant effect of color-coding on retention and recall of gender information. This effect was attributed to the idea that color-coding "[assisted] in the selective attention process of the brain" (Kohler, 2009, p. 17) and drew increased attention to the grammatical information allowing the processing necessary for the information to be moved into long-term memory. Kohler directly cites Sharwood Smith's (1991)

theories of "input enhancement" and Schmidt's (1990) theory that views noticing as a critical first step in language acquisition, and she suggests color-coding as an effective technique in aiding the transfer of input to learner intake. As well, color-coding was considered a means of multiple coding of the material and a way to mentally categorize the new vocabulary and grammatical information to make it more easily retrievable (Nyikos, 1987). Kohler outlines that "color-coding helps learners organize and categorize instructional information creating a pattern for better interpretation and adjustment to the task" (Kohler, 2009, p. 56). Nyikos emphasizes that long-term retention of information can only be achieved if the learner processes input in such a way that it is arranged "into a personally meaningful pattern" (Nyikos, 1987, p. 9). Furthermore, the information needs to be stored with enough elaboration that it can easily be recalled again. Color-coding is suggested as a cue by which to enhance, code, and organize information in a meaningful way so later recall can be facilitated.

Especially Nyikos' (1987) use of pictures in her color-coding experiments is interesting. She cites Paivio's theories of dual-coding and emphasizes the benefit of utilizing a link between visual and verbal information and the associated dual pathways of memory (Paivio, 1971). By linking the verbal gender information to visual cues such as colors and pictures, the information is processed at a deeper level and is recalled more easily. In turn, Nyikos (1987) found that the condition in which participants viewed a color-coded line drawing, returned the most significant results showing the greatest retention of gender information. In contrast, when participants were only shown a line drawing and word with no color-coding, gender recall was the weakest. This seems to speak for the importance of directing increased attention specifically to the grammatical gender information if this formal feature is to be learned. While images may help the learner understand the meaning of a new word, a picture alone may not be particularly helpful in

learning specific elements of form and grammar. As gender must be considered an element of form, visual aids such as color-coding may be more effective in their increased focus on specific grammar elements.

This color-coding technique looks to enhance input, direct attention, and provide a visual system by which to code and categorize gender information. These ideas all tie into previously mentioned theories of learning and memory. Color-coding also seems like a more easily applicable technique that can be used for words alone, in context, and in relation to various forms of gender marking. Further, either the teacher or the student can color-code making it a more or less active technique. Nyikos (1987) further suggests the possible value of having students write or highlight words in a specific color or having students imagine words in a color. This technique lacks an obvious semantic link though. While color-coding may provide a pattern and framework for organizing gender, it does not do much in terms of making gender more meaningful to the learner or enabling a link to a more concrete mental concept. This may be a weakness of the technique for some learners struggling to understand grammatical gender, or it may act as a strength by decreasing the complexity of the technique.

Methodology:

This study intends to modify and simplify the experimental techniques outlined above, in order to apply them to a classroom situation in such a way that they require minimal changes to the established teaching techniques. The aim is to discover if simple adjustments to teaching materials can lead to enhanced gender acquisition and recall, or if more extensive techniques that involve greater explicit focus on gender are necessary to see results.

The modified keyword method used by Desrochers et al. (1989) and Desrochers et al. (1991) seems complex and too involved to be quick and easy for students to apply in a communicative situation. Further, it was found in their experiments that the effectiveness of the technique was significantly reduced when the students were not provided with a keyword. The task of coming up with associative keywords for all vocabulary to be learned may be a great additional load for many teachers and coming up with such many-faceted mental imagery may be too great a load for many students. This study looks to examine if simply the visual association of gendered actors in relation to target word imagery can effectively facilitate gender recall. In keeping with the rationale behind the modified keyword method, this technique still aims to use visual aids to make grammatical gender more concrete and to create a semantic basis to which to link the grammatical information to the target word meaning.

The color-coding technique examined by Nyikos (1987) and Kohler (2009) seems like a far simpler approach to enhancing input and facilitating uptake of gender information. In keeping with the aim to make simple modifications to aid learning, this technique has value for this study. As the established teaching technique used in the classes examined already makes use of pictures during the initial learning phase (the 5-step teaching technique is outlined in more detail below), color-coding learning materials seems like an easy modification that according to Nyikos' (1987) results, might have a significant effect on learning gender as opposed to using pictures and words alone. This technique requires little extra effort on part of the teacher or student, but may draw extra attention to key grammatical information and help facilitate gender acquisition.

Research Questions:

This study intended to compare the gender acquisition of beginning learners of German under experimental conditions utilizing modified gendered actor and color-coding teaching techniques. The experiment conducted aimed to seek answers to several research questions: Do different types of visual enhancement (color-coding, gendered actors) have different effects on students' immediate learning of grammatical gender? Do different types of visual enhancement (color-coding, gendered actors) have different effects on students' retention regarding learning of grammatical gender over time? Is the learning of grammatical gender more difficult for some words than others (immediately and over time), and is this influenced by different types of visual enhancement (color-coding, gendered actors)?

Participants:

A total of 62 participants took part in this study. All were beginning learners of German enrolled in second semester German classes at the University of Alberta. All participants had at least one semester of previous German language learning experience or the equivalent thereof, but did not possess language skills past the basic beginner level. The students were divided into three groups based on their enrollment in a particular course section prior to the beginning of the semester. In this respect, this study relied on convenience sampling, as a form of non-probability sampling, based on criteria of language ability and class enrollment (Mackey and Gass, 2012, p. 81). In order to avoid the possibly confounding variable of teacher effect, the three classes were chosen based on the fact that all three are regularly taught by the same instructor. Each class was presented a different experimental condition. 23 students took part in the control condition, 18

took part in the color coding condition, and 21 took part in the gendered actor condition. Each class period consisted of 50 minutes during which the entire experimental procedure took place.

Materials:

The target vocabulary words fit into the category "gifts." This topic was the next to be discussed in the course textbook at the time of the study. In this way, the study fit into the natural progression of the course and the vocabulary words were at a skill-appropriate level. All words were unfamiliar to the students in the sense that they had not explicitly appeared in the textbook or coursework so far. Words were chosen so that there were an approximately even number of words for each of the three genders and so that there were both morphosyntactically regular and irregular words (i.e. *die* words that did and did not end in *e*), as well as simple and compound words, and all words were concrete nouns. Some words were chosen directly out of the textbook, and the remaining ones were chosen with the above listed criteria in mind for a total of 12 words.

Teaching materials were based on the standard materials and procedures of the 5-step vocabulary teaching technique[1] typically used in the course. This technique involves presenting each vocabulary word as text next to a color picture depicting it, and each word is additionally assigned a number. PowerPoint slides then allow the instructor to present the words and ask students to receptively recall vocabulary by identifying the number of the word in question. After students repeat the words in chorus, slides are presented in which the written-out vocabulary words have been removed, and only the pictures and numbers remain. In a fourth

[1] For a detailed description of the 5-step teaching technique see: Tschirner, E., Nikolai, B., & Terrell, T. (2013). *Kontakte: A Communicative Approach (7th ed.).* New York, NY: McGraw-Hill.

step, students are then asked to productively recall the vocabulary by identifying the word of a numbered picture in question. Lastly, a personalization phase follows that is linked to a topic specific activity. For this experiment, three sets of PowerPoint slides were created (see Appendix A). In the control condition, vocabulary words appeared written in black text, numbered 1 through 12, and appeared with a color picture depicting it (Figure A1). In the color coding condition, only the text was manipulated and appeared in pink for feminine words, in blue for masculine words, and in green for neuter words (Figure A3). In the gendered actor condition, in addition to black text and the picture depicting the noun, as in the control condition, a silhouette of a man, a woman, or a baby appeared in accordance with the gender of the word (Figure A5). Additional slides were made in which the text was taken away and only the numbers and pictures remained (Figures A2, A4, A6 respectively). The color coding and gendered actors were kept constant in these additional slides.

Procedures:

Participants were tested class by class within their regular class time. The instructor was asked to teach each class as she normally would and to teach each of the three classes exactly the same. The only thing that was different in each condition was the presentation of the vocabulary on the PowerPoint slides. The students were first introduced to the topic "gifts" and were asked to think about gifts they received at Christmas. The instructor then presented the new vocabulary and went through the standard 5-step vocabulary teaching technique. First, the slide with text and pictures was displayed, and the instructor presented the words within a linked context describing how she gave each item as a gift to a family member at Christmas. After going through all of the words once, the instructor asked students to identify the number next to each of

the words in the receptive recall phase. The teacher then read each word aloud and the students repeated the word back in chorus. Then the second slide with numbers but without words was displayed, and the students were asked to say aloud each of the vocabulary words in the productive recall phase. Each time, the students were also to include the gender article of the word. After this, the students were asked to get into partners for the final personalization phase. They were all told they won $500 in the lottery and were to think about which gifts they would buy and give their own family members. After a few minutes of time in groups, the students were asked to share some of their sentences with the class. The class was conducted almost entirely in German except for some clarifications and questions. Approximately 30 minutes were spent in each class learning and working with the vocabulary. While the instructor devoted approximately the same amount of attention to each word during her instruction, there was some variance in student use of the new words during the personalization phase with some words being used and repeated more than others. Approximately the last 10 minutes of each class were left for the assessment task. Participants each received a post-learning activity (explained in more detail below) and were asked to fill it out to the best of their ability.

After class, the instructor posted the vocabulary words and pictures on the course webpage where all students had access to them. The following class period for each course section was further devoted strictly to working with the new vocabulary. After this, coursework moved on and it was left up to the students to continue to practice and use the vocabulary words with no more in-class time devoted entirely to this purpose although some of the experimental words came up at random. Reiterating, only some of the experimental target words were from the textbook. Those words that were chosen by the researcher supplementary to the textbook

vocabulary were still available on the course webpage but may have came up less often in class work and homework following initial instruction and may have had less exposure and practice.

Data Collection and Analysis:

The assessment materials consisted of short post-learning activities. Each of the 12 vocabulary words appeared with a blank next to it in a quiz format and the students were asked to fill in *der, die,* or *das* accordingly (see Appendix B). Students were only asked to recall the noun gender not the word itself as well. For the color-coding and gendered actor conditions, students were additionally asked to mark whether they found the slight modification of the vocabulary presentation beneficial, distracting, or if there were other comments.

Four weeks after the initial learning and testing phase, the researcher went back to each of the classes to conduct a delayed post-test. The students were not made aware ahead of time that this would be happening and therefore did not have any warning to prepare explicitly for the quiz. The same post-learning activities were distributed again to all participants, and they were each instructed again to fill it in to the best of their abilities. This delayed post-test was conducted at the end of the course unit the day before the unit exam on which vocabulary from the topic "gifts" was tested along with other vocabulary and grammar. In this respect, students were tested on material they had practiced and should have felt comfortable with at that point.

Each post-learning quiz was scored out of 12. One point was awarded for each correct gender written down and no points were awarded for incorrect answers. Students could receive a maximum of 12 points.

Results:

The resulting test scores did not show large differences between conditions or between the immediate and delayed post-test. The mean scores (out of 12) and standard deviations for the immediate post-test (Quiz 1) were $M = 9.00$, $SD = 1.88$ for the control condition; $M = 8.89$, $SD = 1.57$ for the color-coding condition; and $M = 8.29$, $SD = 2.61$ for the gendered actors condition. The mean scores (out of 12) and standard deviations for the delayed post-test (Quiz 2) were $M = 8.20$, $SD = 1.82$ for the control condition; $M = 9.29$, $SD = 1.83$ for the color-coding condition; and $M = 7.90$, $SD = 1.83$ for the gendered actors condition (see also Table 1 below). Subject test scores were subjected to a repeated measures of analysis of variance (ANOVA) for both Quiz 1 and Quiz 2, with the three independent variables being the three tested conditions (Control, Color-coding, and Gendered actors), to test if condition as a main effect had any significant effect on mean test scores (see Appendix C, Tables C1 and C2). ANOVA results revealed no statistical difference between conditions for Quiz 1, $F(2, 59) = .72$, $p = .49$, or for Quiz 2, $F(2, 54) = 2.90$, $p = .06$ with $\alpha = .05$. However, the results for Quiz 2 indicate scores approaching significance as a result of condition. Post-hoc comparisons using the Tukey Honestly Significant Difference test indicated that no conditions were statistically different from each other in Quiz 1 ($p > .05$ for all comparisons) (see Appendix C, Table C3). Post-hoc Tukey HSD results for Quiz 2 were: Color – Actor, $p = .06$; Control – Actor, $p = .86$; Control – Color, $p = .17$ (see Appendix C, Table C4). While there was again no statistical difference, a comparison of the color-coding and gendered actor conditions, showed results approaching significance with the color-coding condition having the overall highest mean scores ($M = 9.29$, $SD = 1.83$) and the actor condition having the overall lowest mean scores ($M = 7.90$, $SD = 1.83$). A paired sample t-test was used to compare individual test scores between Quiz 1 and Quiz 2 within each

condition to look for significant improvement of scores (see Appendix C, Tables C5, C6, C7). No statistical difference was found for the color-coding condition, t (16) = -.97, p = .35, or the gendered actor condition, t (19) = .90, p = .38 with α = .05. However a statistical difference was found for the control condition, t (19) = 2.22, p = .04. Mean scores for the control condition show a decrease in scores between Quiz 1 and Quiz 2 (M = 9.00 for Quiz 1, M = 8.20 for Quiz 2). This indicates that test subjects in the control condition forgot a statistically significant amount of the tested grammatical gender information between the immediate and delayed post-test. As the t-test did not reveal a significant decline in mean test scores for the color-coding or gendered actor conditions, there is reason to think that these teaching techniques improved retention of gender information, although this should be tested with a larger sample size to increase the power of the t-test. These results also revealed that the color-coding condition was the only condition that saw an improvement of mean test scores between Quiz 1 and Quiz 2 (M = 8.89 for Quiz 1, M = 9.29 for Quiz 2). While this increase was not statistically significant, the post-hoc Tukey HSD results along with the t-test results lend reason to believe that there may be a trend indicating the color-coding condition to be the most effective. Due to the relatively small sample size of the experiment though, more testing is necessary to corroborate these conclusions. Table 1 below summarizes the mean test scores and shows the improvement or decreases in test scores for each condition.

Table 1
Mean test scores out of 12 by condition and quiz

mean test scores	**Control**	**Color-Coding**	**Gendered Actors**
Quiz 1	9.00	8.89	8.29
Quiz 2	8.20*	9.29	7.90

Note. *p<.05 – significant decrease in mean scores

Boxplots show a graphical summary of the experimental test scores by condition (Figures 1 and 2 below). It is apparent that the Quiz 1 scores were similar across conditions with the median lines being at nearly equal levels. The gendered actor condition displays a larger range in scores though. The Quiz 2 boxplot displays an increased effect of condition on mean test scores. The increase in Quiz 2 test scores for the color condition can be seen.

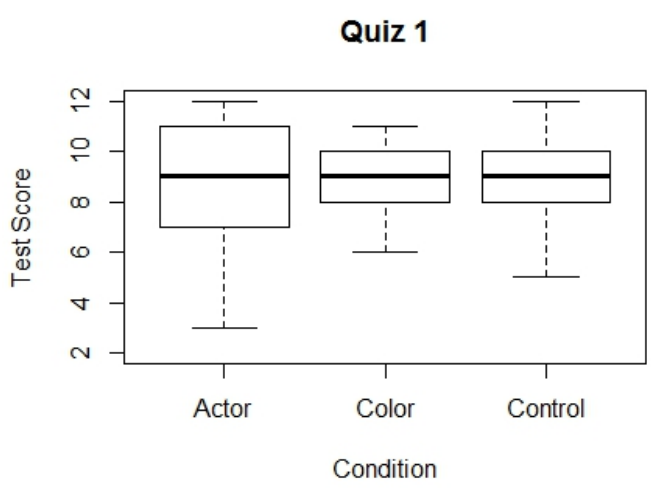

Figure 1: Boxplot of immediate post-test scores

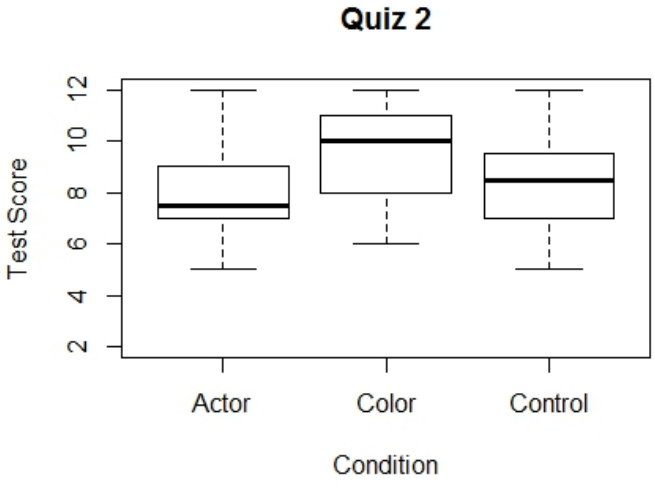

Figure 2: Boxplot of delayed post-test scores

An Item analysis for the error rates per word for each of the quizzes and conditions was also subjected to statistical analysis. A repeated measures ANOVA was conducted for Quiz 1 and Quiz 2 with the 12 experimental target words as independent variables. A significant difference was found for both Quiz 1, $F\,(11, 24) = 4.9$, $p < .001$, and Quiz 2, $F\,(11, 24) = 4.17$, $p = .002$, revealing a clear main effect of word on mean scores. A regression analysis for word error rates revealed an $R^2 = .55$ for Quiz 1 and an $R^2 = .50$ for Quiz 2 indicating that 55% of the variability in test scores for Quiz 1 and 50% of the variability in test scores for Quiz 2 can be accounted for by the variability in the words considered all together. A summary of the error

rates per word for each Quiz 1 and Quiz 2 can be seen in Figures 3 and 4 below. It is evident that certain words were easier to learn and recall the gender for than others. On the immediate post-test especially *die Konzertkarte* and *der Wein* had low error rates while *der Regenschirm* and *das Kuscheltier* seemed hard. On the delayed post-test some changes in terms of difficulty are evident. The words *die Badehose, das Parfüm,* and *der Bikini* all show improvement while *der Kuchen* seemed to pose more difficulties. Overall it is also clear that some words showed similar results across all three conditions while other words like *die Stricknadel* and *das Brettspiel* varied in error rates between conditions. These results indicate a necessary further consideration of why some words appeared to be easier to learn the grammatical gender for than others and what effect, if any, the experimental conditions had on these varied results.

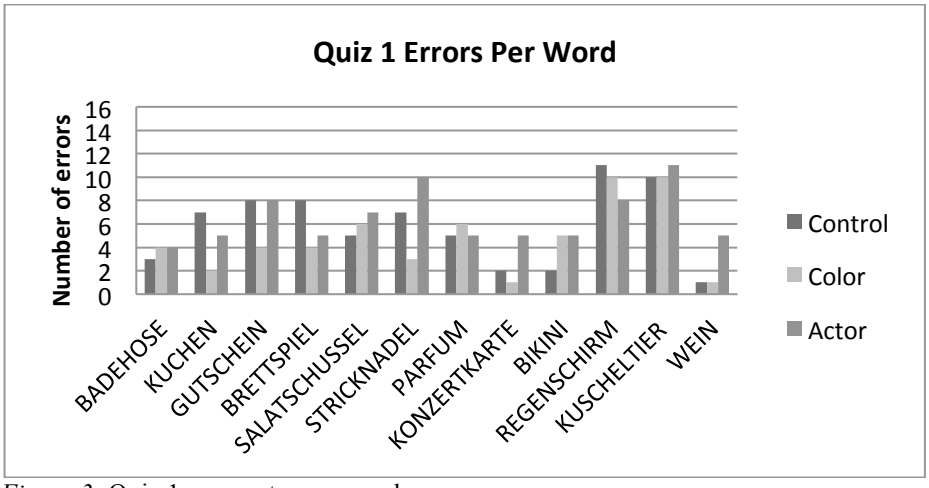

Figure 3: Quiz 1 error-rates per word

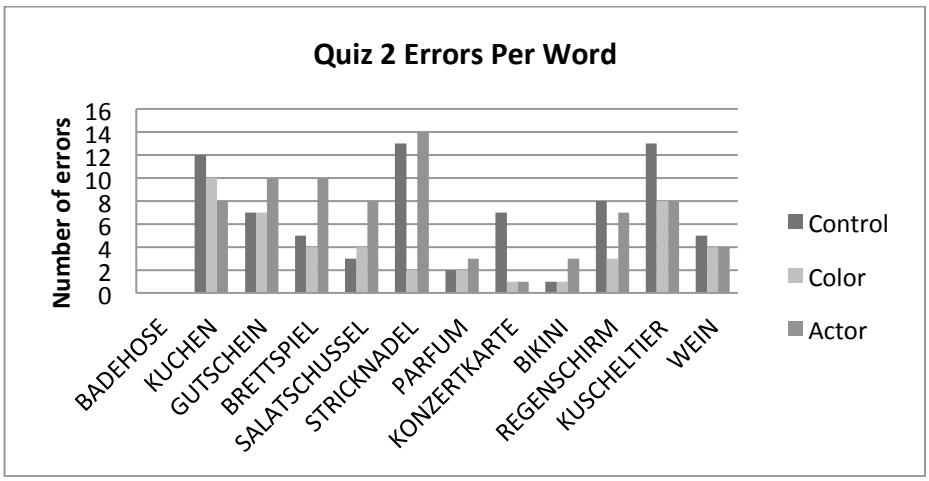

Figure 4: Quiz 2 error-rates per word

Discussion:

This study aimed to compare three different techniques for teaching the grammatical gender of German nouns. The experiment looked to see if simple modifications to the teaching materials of standard classroom teaching practices, based on ideas of visual enhancement, had an effect on the initial learning and retention over time of grammatical gender information among beginning learners of German. The experimental techniques aimed to increase the salience and meaning of grammatical gender via color-coding and gendered actors to improve the attentive processing of this element of vocabulary and to encourage the intake of noun and gender as one unit. The effectiveness of these approaches is discussed below.

Summary of findings:

The results of the initial post-test did not reveal any significant results with the mean scores being very similar across all three conditions. The delayed post-test though, revealed results that displayed more variance. Overall, the effect of the experimental conditions on the quiz scores was shown to be approaching a statistical difference, while not quite reaching

significance. The teaching techniques did have an effect on quiz scores and the learning of gender information, but overall these effects seemed to be minimal and mostly apparent only in a delayed test.

Initial learning of gender information was relatively high across all three conditions and seemed to be affected little by the experimental techniques. Visual enhancements of both types tested had a minimal effect on the short-term learning of gender. A more significant result of the experimental techniques was seen in terms of retention of gender information over time. A significant difference was found between the quiz scores of the immediate and delayed post-test in the control group. As can be seen by the mean scores, the control group scores decreased the most. This seems to indicate that test subjects in the control condition forgot a significant amount of the tested grammatical gender information between the immediate and delayed post-test. In comparison, the mean test scores show that the gendered actor group showed only a slight decrease in their test scores between the immediate and delayed post-test and that the color-coding group actually displayed a slight increase in scores. These changes in mean test scores for the gendered actor and color-coding conditions were not significant though. However, considering that the initial test scores across conditions were fairly high, relatively little change in test scores in these conditions four weeks later lends reason to believe that the experimental teaching techniques improved retention of gender information in comparison to the regular classroom teaching technique alone. While initial learning was not significantly affected by visual enhancements, color-coding and gendered actors did seem to have an effect on gender learning over time.

A comparison of the error rates for each of the words between the quizzes also led to interesting results. A statistical analysis showed a clear effect of word on error rates. This

indicates that across conditions, some words seemed significantly easier to learn than other words. As all words were unfamiliar to students and were chosen based on the same criteria of topic, but differed in terms of length, regularity, and gender, a further analysis is necessary to interpret what led to increased difficulty in gender learning.

The immediate post-test showed that there were certain words that seemed generally harder to learn than other words across all three experimental groups, while some words seemed significantly more or less easy to learn for only particular groups. These results are even more apparent in the delayed post-test results. This seems to indicate that while there are always the same three genders, learning gender does not pose an equal challenge for every word. Examples of this can be seen on a word-by-word basis (see also Figures 3 and 4 above). *Die Badehose* resulted in not a single error across all three conditions on the delayed post-test, and *die Konzertkarte* seemed relatively easy as well in at least the two experimental conditions. Both of these words end in –e and follow the fairly consistent pattern that makes such words feminine (a pattern students have been made aware of in previous instruction), and may have acted as a clue to participants. However, these were the only words tested that follow more consistent gender assignment patterns, and it is therefore difficult to draw conclusions about the benefits of teaching gender "rules." Furthermore, the fact that the control group still seemed to have difficulty with *die Konzertkarte,* despite knowing the rule, may instead be evidence of the improved effects of the color-coding and gendered actor techniques. *Die Badehose* is also a compound word. In German, compound words always take the gender of the last noun component. *Die Hose* is a word students previously learned in the course, and possibly, it was this connection, which influenced the high scores on this word across conditions. A number of other words tested were also compound words. Of them, *das Brettspiel* and *das Kuscheltier*

were ones students also knew the root word of (*das Spiel* and *das Tier* respectively). However, these words did not show consistently better results. Rather, longer words like *das Kuscheltier, der Regenschirm,* and *der Gutschein*, all of which are compounds, seemed to be hard. Comparatively, short words like *der Bikini* and *das Parfüm* seemed to be relatively easy. Interestingly, the word *der Regenschirm* was one of the hardest words across all three conditions on the initial post-test. This word showed significant improvement on the delayed post-test, especially in the color-coding group. The course instructor noted, that this word had come up several times throughout the instructional unit and may have been repeated and practiced relatively more than some other words, which likely affected the improved knowledge of the word's gender. The word *der Kuchen* displayed high error rates on the delayed post-test despite being a short word. This word may have been hard as the students were simultaneously learning the very similar word *die Küche,* which has a different gender, and may have led to confusion. *Die Stricknadel* was a particularly interesting word, as it seemed to be one of the hardest words in the control and actor groups, but one of the easiest for the color-coding group in both tests, with enhanced discrepancy on the delayed post-test. Why this occurred is unclear, and attempts should be made to replicate these results to see if they are truly a result of visual enhancements or if they were particular to this group of participants. Overall, the differences in error rates for each of the words can be attempted to be accounted for by a variety of explanations. However, none of these explanations seem to be overarching or apply to all words of a particular kind. Furthermore, there does not seem to be any convincing evidence that a particular form of input enhancement is more effective for some words than others. In respect to words that may be considered to be more culturally gendered such as *die Badehose* or *der Bikini* that have a contrasting grammatical gender, or *die Stricknadel* that has a comparable grammatical gender, no

significant effect is evident that shows that these words were any more or less difficult to learn with the potential aid of a gendered actor next to them. Variance in error rates between words seems general and is likely inevitable when considering a variety of different words and an overall small sample of words from a language.

In addition, it is significant to note, that on the whole, no clear patterns were evident showing that words of one gender were easier to learn than of another gender. Both among the overall easiest and overall hardest words, each of the three genders are represented. Further, the set of 12 target words was comprised of words appearing in the course textbook as well as ones selected in supplement by the researcher. *Die Badehose, die Salatschüssel, das Parfüm, der Bikini,* and *der Regenschirm* were each words that appeared in the textbook; *der Kuchen, der Gutschein, das Brettspiel, die Stricknadel, die Konzertkarte, das Kuscheltier,* and der *Wein* were each words that only appeared in the experimental teaching materials, which were made available to students online but did not appear in the textbook. Looking at Figures 3 and 4 above, it can be seen that no clear trends are apparent that indicate an effect of vocabulary words that did and did not appear in the textbook. This is significant in terms of independent practice of vocabulary, showing that learning was not significantly affected by the absence of words from one mode of student practice. Some words clearly provided students with greater difficulty, but the data does not show that difficulties were linked to factors of gender group, appearance, or experimental condition.

Discussion of findings:

An interpretation of these results can be attempted in terms of each of the experimental teaching techniques and the theoretical background considerations examined above.

The results of the color-coding group in terms of the increase in scores between the initial and delayed post-test were especially interesting. Students did not engage in any explicit instructed practice of the experimental target words after the initial learning activities and the color-coding group did not receive any instruction or practice significantly different from the other experimental groups. Furthermore, the students were not made aware ahead of time of the delayed post-test and did not have an opportunity to prepare specifically for the quiz. In these respects, a slight decrease in test scores was to be expected in comparison to the immediate post-test. While the increase in scores was not significant, the color-coding group was the only group to display improved results lending reason to believe that this technique might be the most effective of the three in this study.

The color-coding technique was a simple way to enhance the gender information of each noun and visually categorize the words as they were presented. The technique still treated the gender and noun as one unit, color-coding the whole word unit, but attempted to bring extra attention to the grammatical information of the word. It did this in a visually bright way, yet remained simple and possibly less distracting than other techniques. By color-coding, the learning materials focused on enhancing one particular aspect of input and attempted to increase salience and direct student attention specifically to gender. The technique aimed not to overwhelm students with an abundance of new information or new ways of approaching vocabulary learning. Rather, the teaching technique intended to direct attention while not drawing a significant amount of attention away from learning the target words. Target word and

gender were to be learned as one unit with input enhancement supporting the gender component of the word unit and providing an added component by which to code and categorize information to aid later recall. These factors may have helped students recall the information later on as is reflected by the overall high quiz scores.

A further consideration is the socially symbolic colors chosen with which to color-code. While this experiment simply carried over the same colors used in previous research on color-coding techniques (Nyikos, 1987), these colors are not arbitrary. In western cultural conventions, pink is often associated with girls, blue with boys, and green remains a fairly gender neutral color, as is noted by Nyikos (1987, p. 68). While this technique less explicitly links grammatical gender to societal gender, in comparison to the gendered actor technique, such a link can still be interpreted. Students may have made this connection as well. In such a way, the color-coding may have done more than just flag and visually categorize words by their gender, but the colors may have also provided a familiar basis to which to link the new concept of grammatical gender. Not only could words be visually separated out based on their gender, but the colors may also have triggered previous conceptions of gender categorization and have helped students mentally group words as they were being learned – a further factor that may have improved recall.

Turning to the gendered actor condition, it is evident that this technique resulted in the overall lowest scores and showed the greatest range in scores especially on the immediate post-test with some students performing very well under the condition and others very poorly. This may be an indication that this technique was unclear to participants. While some students may have made the connection between the images of the gendered silhouettes and the gender of the nouns, others may have been confused by this technique. The 5-step technique involves

presenting new vocabulary words in an associated context. As the words all belonged to the category "gifts," the presentation of the vocabulary involved conversation about giving gifts to various people. Thus, along with the gender of the nouns, the gender of various actors receiving gifts also appeared in the input students heard. Some participants may have thought that the silhouettes were somehow related to this "gift giving" context rather than linking the actors to the nouns gender, leading to confusion on the post-test later. Furthermore, a qualitative comparison of the three sets of teaching materials shows that the gendered actor teaching materials were the most busy with the slides containing twice as many images and a great deal of information being presented to the students at once. This technique aimed to simplify the mnemonic techniques suggested by Desrochers et al. (1989) by focusing just on the gendered actor component of the technique. However, this approach may still have been too complex and confusing to students. While the extra pictures were intended to help flag gender information and draw attention to that aspect of each of the new words, the black silhouettes may have been less visually clear and may have drawn attention away from the word learning task. Possibly, the ability to visually group and categorize words with the help of the extra images was not as easy as intended, and if participants did not make the connection between the actor and the noun gender, the silhouettes provided no benefits. Rather, then the slides may have seemed only overwhelming and confusing and may have distracted learning rather than enhancing it.

However, it is important to note that students in this group still forgot relatively less in comparison to the control group with their test scores only decreasing slightly. Possibly, the extra images did still lead the students to focus more on the input and consider the materials presented to them in a different way than usual, resulting in more attentive processing of the input. This may have led to better recall later and improved retention in comparison to learning

without any input enhancement. Further, the large range in scores also seems to indicate that not all students were confused and overwhelmed by the technique. Some students performed well on the post-tests and may have made the intended mental associations and focused to a greater degree on gender information and thereby benefitted from the input enhancement via gendered actor pictures. This suggests a need to make the teaching techniques more clear to students to improve their effectiveness.

Lastly, the control condition showed reduced retention of gender information displaying the greatest decrease in test scores over time. This seems to indicate that without any specific enhancement of grammatical gender information in learning input, this technique was the least effective over time of the three techniques tested. However, it is not to be disregarded that initially students in this condition displayed a fairly low error rate in their learning of gender. The 5-step technique used regularly in classroom instruction does not focus explicitly on gender in any of the steps, but overall the students do seem to focus enough on gender information and do seem to consider it part of the noun unit that is to be learned. Initially, techniques employing visual enhancements were equally effective, and while there may be an indication that over time input enhancement can be beneficial, the results of this experiment do not discount the validity of the 5-step teaching technique used.

A further noteworthy result was that students in the color-coding and gendered actor conditions were asked to note on the bottom of their quiz if they found the condition beneficial, distracting, or if they had other comments. These qualitative assessments of the techniques by the participants were particularly interesting for the immediate post-test (the delayed post-test did not involve another interaction with the learning materials and many participants made no further comments). In both conditions, a large number of students noted that they had found the

technique to be helpful. In the color-coding condition, 39% of participants marked the technique as beneficial and in the gendered actor group 67% of participants felt this way. However, in the color-coding condition 61% of students noted that they had not noticed a difference in the learning materials or a difference in their learning. In the gendered actor condition, 24% of students did not notice any differences and two participants noted that they had found the modifications distracting. This shows that the gendered actor materials with the added images seemed to raise more awareness, though this was not always deemed beneficial by the students and was not necessarily reflected in the actual test scores. The color-coding condition may have resulted in positive test scores, but according to participants, the technique might have been too simple and too subtle to raise significant awareness and influence learning. As the point of these teaching techniques was to bring extra attention and saliency to the gender information of the nouns, it was somewhat concerning that many students were not consciously aware of the technique at all. The attempted input enhancement did not seem to register with many participants, and a critical disconnect between teacher input and learner intake was apparent.

Two students also made comments noting that they felt that the techniques would have been more effective as an independent study technique (participant in gendered actor group) and that the technique may have been more beneficial if it had been explained prior to the lesson (participant in color-coding group). These comments provide possibilities for future research. If the study were to be conducted again, the possibility of explaining the particular teaching technique to the students more clearly before the lesson may lead to different results. This may not only help draw increased attention to the specific form of input enhancement and the gender information, but might also make particularly the gendered actor technique more clear and might lead to more significant results.

These forms of visual enhancement seem to have the potential to be an effective aid for learning and improve retention, but the application of the techniques in this experiment fell short and clear effects could not be determined. Student learning assessments show a trend in favor of input enhancement, but comments indicate that potentially greater awareness and participation in the techniques might benefit learning more.

Limitations and Future Research:

While the above discussion of the results does seem to display a trend in favour of particularly the color-coding teaching technique, an analysis cannot discount the fact that the results of the experiment are not entirely conclusive. On the whole, the teaching techniques had a minimal effect on learning. Various considerations of the experimental setup might provide explanation of these results.

The two teaching techniques tested in this study were based on the idea of input enhancement. Input presented to students was manipulated and modified in an attempt to increase the saliency of the grammatical gender of each noun and direct student attention to increase the attentive processing and encourage the integration of this important component of the new vocabulary. The techniques aimed to help students turn input into intake and improve learning and retention of a particular aspect of the noun unit. However, while Sharwood Smith advocates for input enhancement as a teaching tool, he also makes clear that "[i]nput enhancement implies only that we can manipulate aspects of the input but make no further assumptions about the consequences of that input on the learner" (Sharwood Smith, 1993, p. 176). Just because a teaching technique attempts to flag a particular aspect of language in input, does not mean that input enhancement will actually register with the learner and have an effect

on how much and what aspects of input are being processed. The participants' comments on their quizzes make this clear – many students were not aware the teaching materials had been manipulated until they were asked to reflect upon those modifications afterwards. Input enhancement by the teacher did not seem to raise sufficient awareness in many cases to affect learning.

Furthermore, even if a student does notice a change in some aspect of input, it does not mean that this noticing will have the intended effect in the learners mind. Sharwood Smith explains: "Here the operations of the teacher on the input, creating salience from outside, are represented as not necessarily creating the desired salience on the inside, i.e., in the learner's mind. [...] In other words, although learners may notice the signals, the input may nevertheless be non-salient to their learning mechanisms and hence will have no effect on development." (Sharwood Smith, 1991, p. 121). This may also be reflected in a disconnect between student comments and actual quiz scores. While especially in the gendered actor condition, many students marked that they had noticed the modifications and had felt that they were helpful, this was not always reflected in their learning assessments. As well, some students may in fact have noticed the gendered actors but made incorrect connections assuming the pictures were linked to gift giving rather than the intended grammatical gender. This noticing of modifications may actually have hindered learning. The added images on the teaching slides in the actor condition were more apparent and harder to miss, but noticing the changes alone is less significant. Rather the aim was for students not only to notice the input enhancement but also to act upon it to aid their learning – only this can be considered effective input enhancement. The overall quiz scores seem to indicate that this did not always occur in the experiment, and that experimental treatments and learner mental states did not always coincide in the desired way.

This experiment chose not to explain the tested teaching techniques to students before the lesson. In doing so, an attempt was made to avoid a novelty effect that may have resulted from students paying excessive attention to gender information because of the experimental context. The intent was that the techniques be clear enough on their own to affect student learning. Nyikos outlines that "a mnemonic must be transparent enough to allow the unaided information to function in a meaningful way. The mnemonic must support – not overwhelm – the targeted information" (Nyikos, 1987, p. 71). With this in mind, the teaching techniques were kept simple and clear; however, this might not have transferred to participants in the intended way. The associations of the modifications to the gender of each noun may not have been made by the students, and particularly the gendered actor technique may in fact have overwhelmed rather than supported the target information. This consideration calls for further adjustments to the techniques and to the application of them to improve student understanding of them and increase their effectiveness.

A further important consideration for the effectiveness of these teaching techniques goes back to the above outlined idea of attention as a limited mental capacity. Our mind cannot focus on everything at once, so attention acts as a filter and selects what aspects of environmental input are consciously processed (Schmidt, 1990, p. 136). Especially unfamiliar input requires increased focus and little attentive capacity remains for other processing. As skills or aspects of input become more familiar with repeated exposure and practice, they slowly become automatic and require less attention to be processed (Schmidt, 1990, p.136). This idea is relevant and important to this experiment. The teaching techniques tested aimed to act as supports to processing and help focus attention to the gender aspect of input. The use of colors and mnemonics aimed to decrease the processing load and add mental connections in order to

support the learning of gender. However, it is important to consider that participants in this experiment were exposed to at least three unfamiliar concepts to be processed. Students were presented with 12 new vocabulary words they knew they needed to learn, along with the difficult concept of grammatical gender associated with each of the new words, all accompanied by new modifications to familiar teaching techniques to which they were supposed to pay attention. These elements all likely required a significant amount of attention and it may be out of proportion to assume that students could focus adequately on each aspect. While students were aware that they were part of an experiment, they were not told that specifically gender was being tested nor were the experimental teaching techniques explained. Further, while the teaching materials attempted to steer attention, each student had to decide on their own, as in every learning situation, which aspects of the input received his or her attention and what he or she deemed important in the lesson. It is possible that the allotment of limited processing resources did not fall to the experimental teaching techniques and they therefore did not significantly affect learning. The new techniques were unfamiliar and may have been unclear to students requiring a significant amount of attention to understand and apply. Participants may have deemed the unexpected modifications as less important, and competing demands on a limited processing capacity focused attention instead on the expected instructional form and the vocabulary words they knew they needed to learn. Instead of increasing attention to gender information, students may have tuned out the experimental changes and focused in on just learning the new words according to their own familiar techniques.

In addition, Schmidt outlines that language learners "process input for meaning before processing it for form" (Schmidt, 1993, p. 212). Upon first exposure to new vocabulary words, students may have been more concerned with learning the meaning of the new nouns, and

focusing on grammatical gender elements and unexpected modifications to teaching materials were not made a priority. One participant noted that he or she had not noticed the teaching modifications due to being busy writing down the new vocabulary. Schmidt goes on to say that "in order for learners to process form that is not meaningful, they must be able to process informational content at little or no cost to attentional resources" (p. 212). Beginning learners of a language may be sufficiently challenged with getting the message across and may be less concerned with focusing limited attention on correct grammatical formulations and gender information. While gender is an essential component of noun information, for beginning learners still unfamiliar with the concept of grammatical gender, this information may seem extraneous and may initially receive insufficient attention. This leaves two possibilities: either gender information as an element of form can be made more meaningful to the student, or vocabulary learning must be made easier to reduce the load on attentive processing.

It is precisely this problem that this study attempted to address through input enhancement. However, the difficulty in achieving results reiterates the challenge of teaching and learning grammatical gender. If a psycholinguistic model of language acquisition such as a connectionist approach is taken as a starting point, vocabulary input and learning are the basis from which advanced gender assignment patterns are built up, and direct learning of vocabulary and gender should be a focus. Grammatical gender is inherently linked to a noun and represents a mental unit. As such, a more lexical approach to teaching and learning that focuses on memory based approaches may be more valuable than techniques that focus on learning a rule-based system. However, while gender has essential grammatical meaning and significance, it has no semantic significance to the noun unit, and thus may be deemed less important by a beginning learner struggling to communicate meaning. This means that gender may not receive sufficient

attention to integrate this information into the developing formal language system upon initial processing. The close link between grammar forms and lexical meaning becomes apparent in the case of gender, and an appropriate teaching technique that focuses both on grammar and vocabulary learning at once must be utilized. As outlined earlier, Desrochers et al. (1989) found that gender recall was significantly worsened if the noun translation could not be recalled. This further suggests that noun and gender are learned as a unit and that a teaching technique cannot disregard either element. This study tested only if students could correctly recall the gender of the nouns learned. Whether or not students could recall the actual nouns was not tested. A more accurate measurement of gender learning might be reflected by considering how many noun units, including noun meaning as well as gender, can be recalled. Considering both of these factors in respect to experimental teaching techniques may give a more comprehensive picture of the value of these techniques for second language learning. Making gender meaningful enough as an element of form to attract sufficient attention while still focusing on vocabulary learning as such, requires a careful balance of student focus.

There are other limitations to this study that must be taken into account as well. Though three classes were recruited for the experiment, the study consisted of a relatively small participant group making significant results and general conclusions difficult to draw. In addition, due to time limitations, each of the techniques could only be tested on one set of vocabulary words for each condition and students were only exposed to the experimental techniques once. To corroborate the conclusions drawn, the study should be conducted again with a larger participant group and ideally over a longer time span during which students could become familiar with the techniques and learn to apply them to various sets of words. With multiple applications of the techniques a greater number of words could be learned and

considered in an analysis to better determine the varied difficulty of gender learning. As techniques are practiced and become more familiar to students, they also become more automatic and can be more easily applied leaving more equal processing capacity available for both vocabulary and gender learning. Repeated exposure to and use of these teaching techniques might lead to more significant results than can be observed after just one attempt. This may give a more decisive idea of the difficulty of gender learning and about the true effectiveness of each of the techniques and the ways in which they might best be administered in a classroom.

The above outlined considerations are important for any technique a teacher may apply in a learning situation. While input enhancement techniques may try to close the gap between teacher input and learner intake, there is by no means a guarantee that teaching techniques will have the desired effect on student learning. To increase the likelihood that teaching techniques are in fact effective, care must be taken to make them simple and clear, and it is important to make learners aware of strategies and use them regularly. Nyikos (1987) emphasizes that techniques should become a known pattern for the learner and be applied again and again with relative automaticity to have the greatest effect. With this in mind, the data collected in this study provide an excellent basis for future experiments working with these techniques.

Taking this study as a starting point for further consideration and analysis of these teaching techniques, several potential future research questions can be raised. Firstly, a further study could explain the teaching techniques in question before the lesson. This could help prime students for the study and increase awareness and familiarity of the techniques. Secondly, this experiment looked at two techniques strictly applied from a teaching side. These techniques could also be applied more actively by students having them code and manipulate language input themselves. If these techniques are used simultaneously as both a teaching and learning

technique, there may be an enhanced effect in learning. These questions may shed light on the most effective ways a teacher can guide students through the complex task of learning grammatical gender.

Pedagogical Implications:

If these results can be replicated and improved in further studies, there may be evidence that modifications as simple as color-coding learning materials can aid students in acquiring a persistently difficult aspect of German. Such techniques would be easy to apply in a classroom by both teachers and learners. Rather than asking for an extensive reorganization of classroom practices, these techniques aimed to keep modifications simple and utilized the basis of a proven teaching technique to build from. Grammatical gender is not only a difficult concept for language learners to acquire but is also a difficult concept to teach in its amorphous position between lexical and grammatical element. The teaching techniques suggested above require further evidence and support but may be a way in which teachers could help students take input and turn it into knowledge of their own to apply to language situations in the classroom and beyond. These results could have pedagogical implications for Second Language Education across gender marking languages.

Conclusion:

Der, die, or *das* are three small German words that cause second language learners to scratch their heads in confusion and lead to persistent grammar errors that let a learner be distinguished from a native speaker even into advanced proficiency. The grammatical gender of German nouns poses both learning and teaching challenges. This paper has argued in favour of a

connectionist model of language learning as a basis for the formulation of effective teaching techniques. This approach suggests that beginning language learners focus on learning noun and gender information as one unit stored in the mental lexicon from which patterns of gender assignment can then be built up. To make the ambiguous and unfamiliar concept of grammatical gender more meaningful and salient in language input, different forms of visual input enhancement were suggested. Color-coding and the inclusion of gendered actor images in teaching materials were employed in an attempt to focus student attention so gender and noun information could be turned into intake and learned simultaneously. In such a way, meaning and formal elements of new nouns were to be emphasized giving gender a more tangible form and a way to code and categorize the information to make it more easily retrievable. The results of the experiment showed that short-term learning of grammatical gender was affected little by visual enhancements. On a delayed post-test though, the retention of gender seemed to be more affected by the two forms of visual enhancement tested with color-coding showing the overall most positive results. In accordance with the fact that some words are more and less difficult to learn in a language, the learning of gender also seemed to be unequal across words. However, no conclusive results could be drawn about what aspects of a word make gender learning easier or whether a form of input enhancement is more or less effective for a type of word. Overall, it was made clear that the gap between teacher-formulated input and student intake and later output, is not easily closed. Teaching techniques may attempt to make learning easier and focus on an aspect of language. Ultimately though, learning occurs within the student and can only be influenced from the outside to a limited extent. In the case of the techniques tested here, attention and learning were not affected as successfully as was aimed, and further research is necessary to discover how grammatical gender learning can best be facilitated.

References

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. *American Psychologist, 30*, 821-828.

Delisle, H. H. (1985). The Acquisition of Gender by American Students of German. *The Modern Language Journal, 69*(1), 55-63.

Desrochers, A., Gélinas, C., & Wieland, L. D. (1989). An application of the mnemonic keyword method to the acquisition of German nouns and their grammatical gender. *Journal of Educational Psychology, 81*(1), 25-32. doi:10.1037/0022-0663.81.1.25

Desrochers, A., Wieland, L. D., & Cote, M. (1991). Instructional Effects in the Use of the Mnemonic Keyword Method for Learning German Nouns and Their Grammatical Gender. *Applied Cognitive Psychology, 5*(1), 19-36.

Gass, S. M. (1997). *Input, interaction, and the second language learner.* Mahwah, NJ: Lawrence Erlbaum Associates.

Gass, S. M., & Selinker, L. (2001). *Second language acquisition: an introductory course (2nd Edition).* Mahwah, NJ: Lawrence Erlbaum Associates.

Hoff, E. (2009). *Language Development.* Belmont, CA: Wadsworth, Cengage Learning.

Kohler, T. (2009). *Using neurocognitive theory to develop a new approach for teaching German grammar the effect of color-coding German grammar on language acquisition. Dissertation Abstracts International, A: The Humanities and Social Sciences,*

Kopcke, K., & Zubin, D. A. (1984). Six Principles for Gender Assignment in German: A Contribution to Natural Classification. *Linguistische Berichte, 93*(Oct), 26-50.

Levine, G. S. (1999). Gender Assignment to German Nonsense Nouns: What Does the Native Speaker Know That the Non-Native Speaker Doesn't? *Proceedings of the Annual Boston University Conference on Language Development, 23*(2), 397-406.

Mackey, A., & Gass, S. M. (2012). *Research Methods in Second Language Acquisition.* West Sussex, UK: Blackwell Publishing Ltd.

Menzel, B., & Tamaoka, K. (1995). Der? Die?? Das??? Gender Assignment by Beginners: Chance, Cramming, or Strategy? *Deutsch Als Fremdsprache, 32*(1), 12-22.

Nyikos, M. (1987). *The effect of color and imagery as mnemonic strategies on learning and retention of lexical items in German*. *ProQuest Dissertations and Theses,*

Paivio, A. (1971). *Imagery and Verbal Processes.* New York, NY: Holt, Rinehart and Winston.

Rogers, M. (1987). Learners' Difficulties with Grammatical Gender in German as a Foreign Language. *Applied Linguistics, 8*(1), 48-74.

Salamoura, A. (2007). Grammatical Gender in the Bilingual Lexicon: A Psycholinguistic Approach. In I. Kecskes & L. Albertazzi (Eds.), *Cognitive Aspects of Bilingualism* (pp. 99-151). Dordrecht, The Netherlands: Springer.

Salmons, J. (1993). The Structure of the Lexicon: Evidence from German Gender Assignment. *Studies in Language, 17*(2), 411-435.

Schmidt, R. W. (1993). Awareness and Second Language Acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Schmidt, R. W. (1990). The Role of Consciousness in Second Language Learning. *Applied Linguistics, 11*(2), 129-158.

Sharwood Smith, M. (1991). Speaking to Many Minds: On the Relevance of Different Types of Language Information for the L2 Learner. *Second Language Research, 7*(2), 118-132.

Sharwood Smith, M. (1993). Input Enhancement in Instructed SLA: Theoretical Bases. *Studies in Second Language Acquisition, 15*(2), 165-179.

Tschirner, E., Nikolai, B., & Terrell, T. (2013). *Kontakte: A Communicative Approach (7th ed.).* New York, NY: McGraw-Hill.

Appendices

Appendix A: Teaching slides used to execute the 5-step technique for each condition

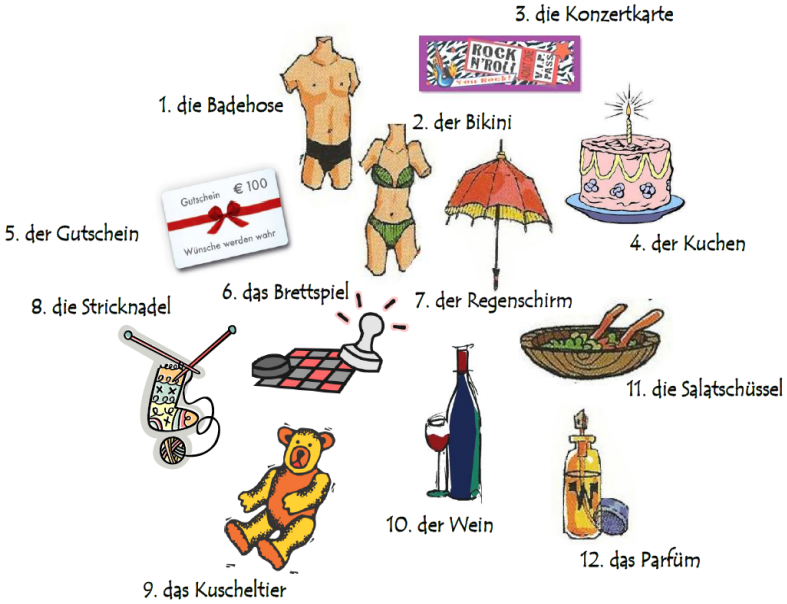

Figure A1: Teaching slide – control condition, with vocabulary

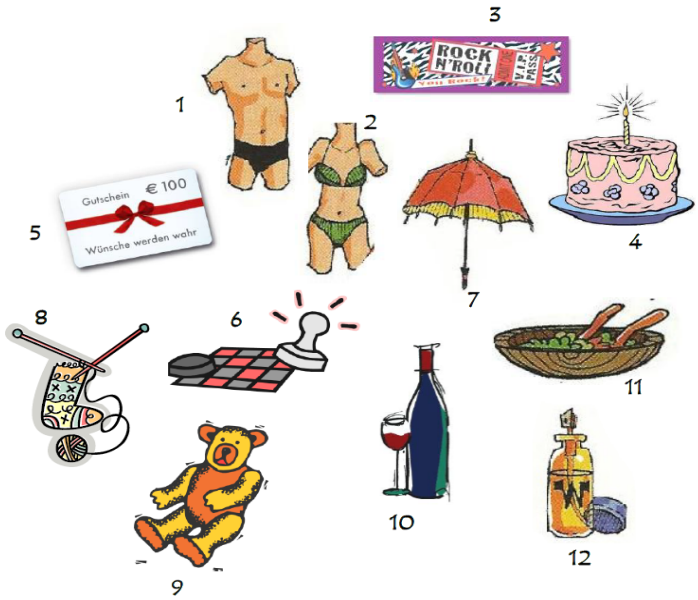

Figure A2: Teaching slide – control condition, no vocabulary

Figure A3: Teaching slide – color-coding condition, with vocabulary

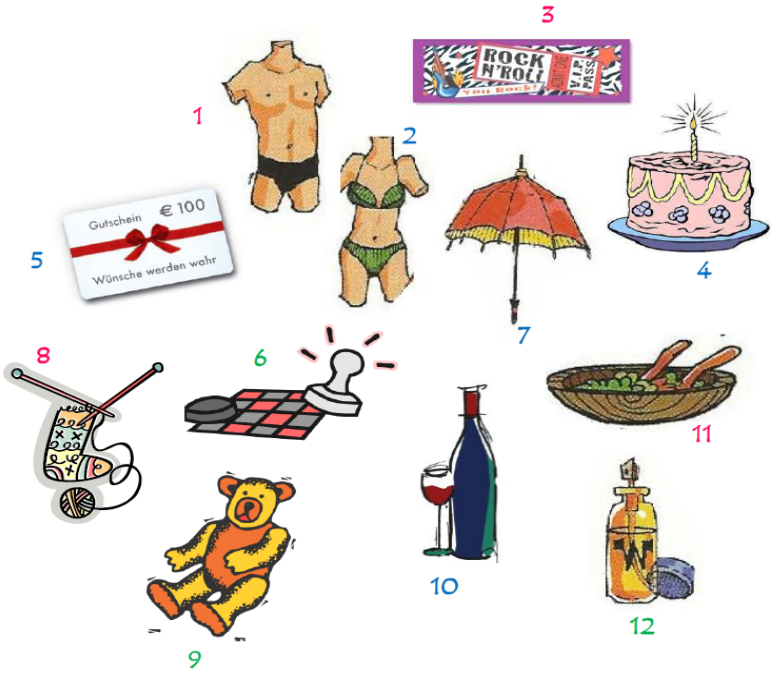

Figure A4: Teaching slide – color-coding condition, no vocabulary

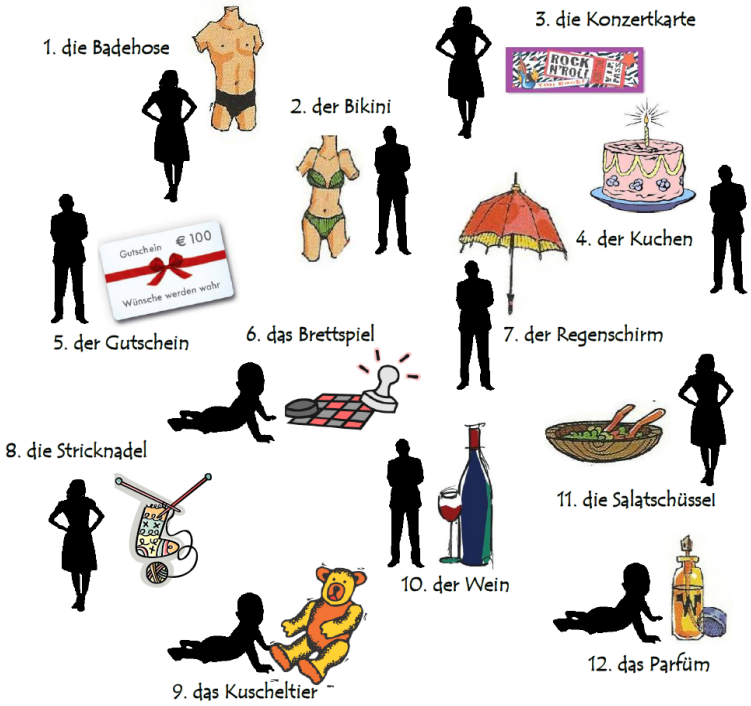

Figure A5: Teaching slide – gendered actor condition, with vocabulary

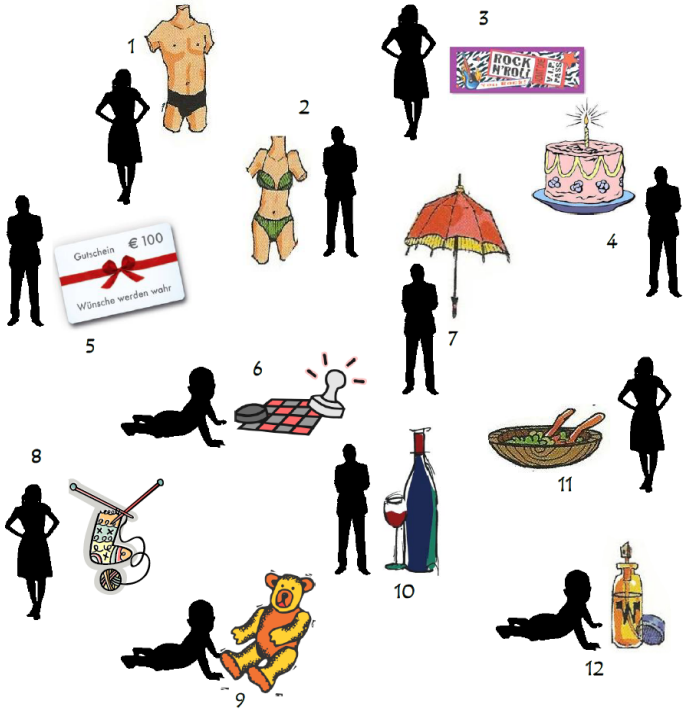

Figure A6: Teaching slide – gendered actor condition, no vocabulary

Appendix B: Sample Post-Learning Activity

Please provide the last 4 digits of your student ID number: _____

Post-Learning Activity

This activity aims to see how much you remember from the in-class vocabulary lesson. Please fill in the definite article (Der, Die, or Das) for each of the vocabulary words listed below to the best of your ability.

_____ Badehose	_____ Parfüm
_____ Kuchen	_____ Konzertkarte
_____ Gutschein	_____ Bikini
_____ Brettspiel	_____ Regenschirm
_____ Salatschüssel	_____ Kuscheltier
_____ Stricknadel	_____ Wein

How did you like the slight modification (color-coding) of the vocabulary presentation?

☐ **Beneficial**

☐ **Distracting**

Other comments:

Thank you for your participation in this study!

Figure B1: Sample quiz used for color-coding condition

Appendix C: Summary of Statistical Results

Table C1

Results for factorial ANOVA: Quiz 1 by Condition

Condition	*Mean (out of 12)*	*SD*	*N*	*p-value**	*F-Statistic*	*df*
Control	9.00	1.88	23	0.49	0.72	Condition=2
Color-coding	8.89	1.57	18			Residuals=59
Gendered actors	8.29	2.61	21			

Note. *α=.05

Table C2

Results for factorial ANOVA: Quiz 2 by Condition

Condition	*Mean (out of 12)*	*SD*	*N*	*p-value**	*F-Statistic*	*df*
Control	8.20	1.82	20	0.06	2.90	Condition=2
Color-coding	9.29	1.83	17			Residuals=54
Gendered actors	7.90	1.83	20			

Note. *α=.05

Table C3

Results for Tukey Honest Significant Differences: Quiz 1 by Condition

Condition	*p-value**
Color - Actor	0.64
Control - Actor	0.50
Control - Color	0.98

Note. *α=.05

Table C4

Results for Tukey Honest Significant Differences: Quiz 2 by Condition

Condition	*p-value**
Color - Actor	0.06
Control - Actor	0.86
Control - Color	0.17

Note. *α=.05

Table C5

Results for paired t-test: Control Condition

Quiz	*Mean (out of 12)*	*SD*	*N*	*p-value**	*t-statistic*	*df*
Quiz 1	9.00	1.88	23	0.04	2.22	19
Quiz 2	8.20	1.82	20			

Note. *α=.05

Table C6

Results for paired t-test: Color-coding Condition

Quiz	*Mean (out of 12)*	*SD*	*N*	*p-value**	*t-statistic*	*df*
Quiz 1	8.89	1.57	18	0.35	-0.97	16
Quiz 2	9.29	1.83	17			

Note. *α=.05

Table C7

Results for paired t-test: Gendered actor Condition

Quiz	*Mean (out of 12)*	*SD*	*N*	*p-value**	*t-statistic*	*df*
Quiz 1	8.29	2.61	21	0.38	0.90	19
Quiz 2	7.90	1.83	20			

Note. *α=.05