# Contents

# 1. Introduction

Gender, as defined by linguistics, has two different aspects: the semantic and the grammatical. In languages with semantic gender, such as English and Chinese, nouns carry only biological sex information; on the other hand, in languages with formal gender, such as Italian, French, and German, all nouns are marked according to their genders, either masculine or feminine or in some languages also neuter. The genders of the nouns do not only carry information concerning referent (Vigliocco & Franck, 1999), but at the same time, also affect mental processes. Recent research suggested that information of different dimensions (i.e. morphologic and phonological) could be activated when identifying nouns. Besides morphologic, phonological, and semantic information, grammatical genders, also referred to as lemma information, are included in these dimensions. Moreover, semantic and phonological information (i.e. semantic similarity and phonological similarity) are also involved in false memory, which involves free recall tasks and recognition tasks. It showed that the more similarities or associated characters that the learned and lure items share, the higher proportions of false memory to lures occurred (Buchanan, Brown, Cabeza & Maitson, 1999; Conway, Gardiner, Perfect, Anderson, & Cohen, 1997; Dewhurst & Farrand, 2004; Dewhurst, 2001; Dewhurst & Anderson, 1999; Dodd & Macleod, 2004; McDermott & Watson, 2001; McDermott, 1996; Roediger & McDermott, 1995; Sommers & Lewis, 1999; Watson, Balota & Roediger III, 2003; etc.). However, as one of the essential activated nodes during identifying words, the gender information was not yet discussed with false memory. It is suggested that the grammatical gender similarity will play a role in

the studies of false memory, according to what was found in the studies of lexical access and word production with gender information.

## 1. 1. Markings of Grammatical Gender

There are at least two kinds of gender-markings relative to the grammatical gender that can influence the lexical access and the word production. In German, for instance, masculine nouns require the masculine form of the definite article *der*, while feminine nouns require the feminine form *die* and neuter nouns require the neuter form *das*. Additionally, grammatical genders also refer to some transparent endings. Examples of this can be seen with words ending in -*tion*, -*ung*, -*heit*, or -*keit* are feminine; words ending in -*chen*, -*lein*, -*sal*, or -*sel* are neuter; words ending in -*ant*, -*ent*, -*ismus*, or -*us* are masculine; words with -*a* ending are mostly, but not always, feminine, etc. Therefore the influence of the grammatical gender in lexical accessing or nouns producing might be caused by two kinds of gender-markings: determiners (i.e. adjectives, definite articles, and indefinite articles), or suffixes of words. Furthermore, some researchers suggested that crucial to the concept of grammatical gender is the agreement in the gender between the noun and other items when comparing the effects from determiners and suffixes (Comrie, 1999).

## 1. 2. Grammatical Gender and Lexical Accessing

### 1. 2. 1. The Influence of the Grammatical Gender from the Suffix in Lexical Accessing

Bates et al. (1995) claim that gender is an inherent property of nouns that can be retrieved at the moment of lexical access for words presented out of context, while the gender assignment for adjectives and other modifiers depends upon the noun that they modify, and is therefore not directly accessed. The factors that influence gender processing in Italian and the influence of gender on lexical access in tasks that do or do not require conscious reflection on this grammatical property are explored in this study. Two tasks were employed for this purpose: repetition of spoken words and a gender-monitoring task. The authors found effects of phonological, but not of semantic gender-marking. Italian participants took less time to classify a word as masculine or feminine if that word ended in a phonologically transparent vowel (*o* for masculine, and *a* for feminine), when compared with words that ended in a phonologically ambiguous vowel (*e* for both masculine and feminine). Participants also made more errors on *e*-final words. The authors suggest that a phonological ending is the cue of the grammatical gender. Italian participants made use of multiple cues to assist with word recognition, permitting them to identify many words before their uniqueness point. Those cues made it easier for participants to recognize words and to track coindexed forms across a complex discourse.

Colé, Pynte, and Andriamamonjy (2003) found a gender bias with the morphological ending *e* in French. That is, there were different lexical decision time

for feminine nouns and masculine nouns. In French, a phonological form can be the marker of the grammatical gender; however, there are always exceptions. These exceptions change the *predictive value* of an ending (PVE). For instance, 42% of nouns ending in *e* are masculine and 58% are feminine, therefore *e* has a low predictive value. The authors investigated whether participants could predict the last letter from the gender of visually presented words (gender-to-ending direction) or could predict the gender of visually presented words from the endings (ending-to-gender direction) with the low PVE ending *e* in visually-presented word recognition tasks. Participants were instructed to decide whether a presented string of letters was a real French word or a pseudo word. All real words in their experiment ended with "*e*". An interaction was found between gender and PVE. The time to make lexical decisions was shorter for masculine items with a high PVE, than for masculine items with a low PVE; whereas PVE exerted little influence for feminine items. It is important to note that the time to make lexical decisions was shorter for feminine nouns than for masculine ones. The authors claim that although *e* is not predictive among nouns ending in *e* in French, it is still more typical for feminine nouns (73% of the feminine nouns end with the letter *e*) than for masculine nouns (67% of the masculine nouns end with a letter other than *e*). In this way, the presence versus the absence of the letter *e* at the end of a given noun is relatively predictable once the gender of that noun has been determined.

This means that people will use multiple cues, which make the task easier, to identify words (Bates et al, 1995). One of these cues is the ending of a word, which is

one of the markings of the grammatical gender. However, exceptions exist where endings of words are gender-ambiguous; such an ending can be the marking of either a feminine or masculine word. It is suggested that grammatical gender can be one of the cues used in recognition tasks only when it can be recognized either with or without transparent endings. Otherwise, phonological similarity or morphological similarity, but not a robust grammatical gender effect, can be concluded on false memory even when the advantages of transparent endings can be found in recognition tasks.

### 1. 2. 2. Influence of Grammatical Gender from the Syntactic Context in Lexical Accessing

Besides transparent endings or nouns, syntactic context, which includes definite/indefinite articles and adjectives, can be a cue in identifying nouns.

Recent research by Bates et al. (1996) discusses the influence of these cues in lexical accessing, using adjectives to prime the following nouns. Whether or not the "conscious awareness" of gender influenced the word, recognition task was a concern. In their research, adjectives and nouns, both of which have either congruent endings (i.e. *a* for feminine and *o* for masculine) or ambiguous endings (i.e. *e*), were presented by native Italian speakers. Subjects were instructed to repeat the nouns presented directly after the adjective as soon as possible in the Word Repeat Task, and make judgments as to the gender of nouns in the Gender Monitoring Task. Robust priming effects were found when target nouns were preceded by a gender-marked adjective

prime. In the word repetition task, facilitation occurred when nouns were primed by a gender-marked, gender-congruent adjective, while inhibition occurred when nouns were primed by gender-ambiguous adjectives; however, in the grammatical monitoring test, only an inhibition effect was found. This concludes that gender priming, with which participants' attention was drawn to the gender marking, regardless of whether or not the task requires metalinguistic awareness of the gender dimension. The explicit attention to gender is not required for priming to occur.

Bentrovato, Devescovi, D'Amico, Wicha, and Bates (2003) used the priming paradigm as well and found a strong facilitation with both gender and semantics more precise in the priming words. They argue that different sources of information are combined on-line to predict, anticipate, or preactivate lexical targets.

However, by investigating the auditory recognition process in German, Bölte and Connine (2004) argue that gender information is not utilized early in the word recognition, though the presence of the gender consistency between article and noun facilitated a subsequent response. Phoneme monitoring, which is very sensitive to the degree of lexical activation, is used in this research. Based on the "similarity effects" discussed by Connine, Titone, Deelman, and Blasko (1997), Bölte et al. (2004) expected a similarity mismatched effect with German materials, because of a set of pseudo words that mismatched a word by a single phoneme was used. The idea of "Similarity effects" claims that derived non-words are similar enough to produce successful activation; lexical effects at the phoneme level take time to accrue; and that fast phoneme monitoring responses occur prior to the sufficient lexical feedback

(Connine et al. ,1997). Thus, Bölte et al. (2004) expected that derived pseudo words, with either very similar mismatched phoneme (Pajonett) or dissimilar mismatched phoneme (Fajonett), were created from each real word (Bajonett). If the presence of a valid article operated to "pre-activate" only those words consistent with that gender, lexical activation could be influenced by gender information. However, although listeners were sensitive to lexical information in the task, no indication showed the valid cue served to mitigate the reduction in the activation due to mismatching segments. The assumption that gender information was not utilized during early spoken word recognition was proven.

Grammatical gender marking plays an essential role in the process of auditory recognition. Combining the auditory recognition process and the Eye-tracking paradigm, Dahan, Swingley, Tanenhaus, and Magnuson (2000) found two factors that affect auditory word recognition: cohort and gender-marking (gender marked by valid articles). In the cohort model, a "cohort" of lexical candidates was activated by the onset of a word. Activated candidates that became inconsistent with subsequent information dropped out of the cohort. Recognition was achieved when one candidate remained in the cohort or when the activation of one candidate was sufficiently greater than that of the other candidates (Dahan et al., 2000). When no gender-marked articles were used as research materials, "cohort" was the only factor that affected word recognition. Fixations of gender-matching competitors were as frequent as that of gender-mismatching competitors. Whereas, when gender marked articles were used, the effect of cohort disappeared. Dahan et al. (2000) claimed that the lexical access

was faster, and the cohort competition was reduced when the target was preceded by a gender-marked article.

### 1. 2. 3. Comparisons between Effects from the Suffix and Syntactic Context in Lexical Accessing

Some researchers focus on comparing effects from the suffix and syntactic context in lexical accessing. Some suffixes are used by either feminine or masculine nouns with fewer exceptions and the gender of words can be more easily identified by these suffixes. Robust regularity effects from suffixes were found only when the dominate influence by the presence of syntactic context was artificially prohibited (Gollan & Frost, 2001, Taft & Meunier, 1998, with Hebrew and with French respectively). Taft et al. (1998) used gender judgment tasks with either bare nouns or place-names (e.g. *Michigan*, *Russie*, and *Bresi*) with article(s). The authors assigned nouns with transparent gender suffixes and fewer exceptions to be "regular nouns", whereas words with the low gender-informative suffix to be "irregular nouns". Two important rules are in French when using the indefinite/definite articles: (1) names of places that are larger than a city take the definite articles, but not the indefinite articles. One can have *le Michigan*, *la Russie*, and *le Bresil*, but not *un Michigan*, une *Russie*, or un *Bresil*; (2) when a place-name begins with a vowel, the definite article itself reduces to *l'* and therefore cannot be used to distinguish masculine and feminine. Information about articles cannot be used to decide the gender of a place-name with vowel onsets (Taft et al., 1998). These place-names also have congruent gender-marked endings. A clear

influence was found from the orthographic information at the end of the word by using bare nouns. When using place-names, the advantage from predictable endings was depressed, making it difficult for French speakers to determine the gender of words of which the articles provided no information (i.e. place-names with a vowel onset) despite that they have highly informative endings. In a further study, place-names were displayed either with a preceding adjective or with a following adjective. The same effect of onset was found as in the previous task. The only form of the article, rather than its form when an adjective intervened between the article and the noun, which was consulted for the determination of gender, was the one that directly attached to the noun.

Different from French, Hebrew has its own gender-marking system. In Hebrew (Gollan et al., 2001), masculine singular nouns are unmarked (e.g., sefer [book]), whereas feminine singular nouns are usually marked by the suffixes /ah/, /et/, or /it/ (e.g., sapah [couch], rakevet [train], zavit [angle]). In addition, masculine nouns are usually pluralized with the masculine plural suffix /im/, and feminine nouns are usually pluralized with the feminine plural suffix /ot/. Certainly there are some exceptions (e.g., feminine words without explicitly marked ending). In Gollan et al.'s research (2001), gender-decision tasks and grammaticality judgments were used with either low proportion (grammatical subjects with masculine, regular feminine and exception feminine nouns in proportions similar to the frequency of their occurrence in Hebrew) or high proportion (all explicitly marked nouns were excluded) of irregular nouns. Thus gender decisions could be based upon the gender-suffix cues when using the low

proportion irregular nouns while not when using the high proportion irregular nouns. Because unmarked nouns are usually masculine, regularity is still a factor in the latter situation. Strong effects of transparency or regularity (Bates et al., 1995) were found with the situation of low proportion of irregular nouns. In other words, despite the differences between Hebrew and French in the gender-marking system, gender processing in Hebrew is similar to French, with respect to the effects of marking on gender decisions. When the proportion of irregular nouns was raised (50%), regularity effects were independent of the presence of explicit morphological markers (i.e., as in feminine suffix /ah/) in the experimental item set. Moreover, in a grammaticality judgment task, the effect of gender-marking regularity on decisions about grammatical pairs was absent. However, although similar to the results of the gender-decision tasks, there is still a clear effect of gender-marking regularity on decisions about ungrammatical pairs. Gollan et al. claims that the detection of correct gender agreement was far less affected by the gender-marking regularity relative to the explicit gender identification. Namely, gender is accessed most efficiently in the presence of syntactic context (Roelofs, 1992). Yet, different processing mechanisms were involved in detecting agreement violations relative to processing correct gender agreements.

### 1. 2. 4. Strong Grammatical Gender Effects from Mother Language

So far, syntactic context, especially definite/indefinite articles, and suffixes have been proven to be essential cues to the grammatical genders, which can influence lexical accessing. However, evidence shows that receiving these cues is not the only way that people identify the gender of words, although they make the grammatical gender effect more remarkable. Some recent studies show an inherent gender effect which depends on the rules of the mother language, despite the given gender cues which belong only to the second language.

In research by Lew-Williams and Fernald (2007), Spanish-learning children (34–42 months old) were tested in an eye-tracking procedure to explore whether they could orient pictures representing the auditory targets with the gender-marked articles. In this research, the speech stimuli were simple Spanish sentences ending in familiar object names (feminine or masculine), with which there was always a gender-informative article (e.g., la for feminine and el for masculine forms of "the"). With each of the sentences, children were given two pictures, one of which was the target object. These two pictures may or may not share the same gender. Lew-Williams and Fernald found that children made a faster identification to pictures that were not of the same gender. The eye-tracking data showed that, at the beginning of a noun phrase, the looking to the target was by chance; then, as the article and noun unfolded, the children began to orient to the target picture as they identified the correct referent. In other words, the young Spanish-learning children identified the referent of a noun more rapidly when the gender-marked article preceding the noun was potentially informative.

These results are comparable to the results found when the subjects were adults. In the same study, it was also robust that the adults made use of the gender-markings on the pronominal articles to facilitate lexical access. This ability was assumed a characteristic only of native adult speakers who learned their gender-marking language as their first language.

Other evidence for gender effect from the mother language was reported by Weber and Paris (2004) with eye-tracking experiments. In their report, both German and French participants received a series of auditory sentences presented in German when the target pictures were displayed. All these target pictures referred to either feminine or masculine German/French nouns. Each target was paired with an onset phonological-related noun (competitor) and a phonological unrelated noun (distractor) in German. The target was always of the same gender in German and French, but the gender of the competitor divided the pairs into two groups (i.e. "same-gender" pairs and "different-gender" pairs, the target and the competitor shared the gender in both languages, or the target and the competitor shared the gender in German, but were of different genders in French). Participants were instructed to click on the object on the screen that was mentioned in an auditory sentence. The appropriate definite article was always displayed with the auditory target. Eye-tracking data showed that French listeners fixated the competitors more than the distractors in the same-gender pairs, but no difference was found between the competitors and the distractors in the different-gender pairs. German listeners fixated longer on competitor objects not only in the same-gender pairs, but also in the different-gender pairs. In different-gender pairs, the

gender information carried by the articles in German could not constrain competitor activation, but the French gender information could. Results support that the grammatically gender-marked effect (definite article) from the non-native language is reduced because of the constraint from the native gender information.

The grammatical gender rules were received when nouns were learned. Bilinguals will only use rules of their mother language during lexical accessing but not the rules of their second language, although transparent cues from determiners were given. The rules, but neither gender-marked endings nor gender-marked determiners, are indeed the factors that influence the lexical access. In addition, gender-marked endings and gender-marked determiners can enhance such a gender effect. This idea makes the hypothesis become possible. It has been proven that the similarity of grammatical gender is one of the important factors, besides phonological similarity, morphological similarity and semantic similarity, which influences false memory. Since the grammatical gender is proven to be activated mainly by using the gender rules of a mother language rather than gender-marked endings (phonological factors), or gender-marked determiners (morphological factors), the influence of grammatical gender to false memory can be discussed.

## 1. 3. The influence of Grammatical Gender when Producing Nouns

The aforementioned studies support the idea that grammatical gender plays an essential role in lexical accessing. Moreover, the following research shows us that grammatical gender could influence word production. This again supports the idea of the importance of grammatical gender effect. In addition, grammatical gender was proven activated individually. With all the above evidence, grammatical gender is available to be discussed with false memory.

### 1. 3. 1. Lemma or not?

On the facet of noun production, given the autonomy of grammatical gender from semantics and phonology, most prominent psycholinguistic models postulate that gender information is stored as a property of nouns at a representational level different from those specifying the corresponding conceptual and phonological information (Caramazza, 1997; Caramazza & Miozzo, 1997; Cubelli, Lotto, Paolieri, Girelli & Job, 2005; Levelt, Roelofs, & Meyer, 1999; Macoir & Béland, 2004; Schriefers & Teruel, 2000; Vigliocco, Vinson, Indefrey, Levelt, & Hellwig, 2004). Several models about producing nouns were proposed (Cubelli et al., 2005). Two models insist on the importance of lemma: two-stage model and WEAVER++ (lemma model). The two-stage model concerns non-verbal conceptual representation maps into modality neutral lemmas at the first stage of lexical selection and lexical forms are retrieved at a separate stage on the basis of an address provided by the lemmas. According to this theory of lexical levels, a lemma is a lexical record that encodes a word's semantic and

grammatical features, but not its phonological or orthographic properties (Badecker, Miozzo, & Zanuttini, 1995). The WEAVER++ (lemma model) theory, a model originally proposed by Roelofs (1992) and refined by Levelt, Roelofs, and Meyer (1999), grammatical gender plays a very essential role in producing nouns. "Lemma model" assumed three cognitive layers in producing nouns: a top layer, which describes the word's meaning by means of a network of conceptual connections; a lexeme layer, which specifies the word's phonological form; and a middle layer, called abstract lexical representation, or lemma, which is connected to nodes representing the word's syntactic properties, such as grammatical gender. In this model, the phonological form of the target word becomes activated only after the corresponding lemma has been selected, which in turn is activated by its corresponding conceptual node. Grammatical gender is an activated node that people can be aware of. Without such awareness, noun-phrases cannot be produced successfully.

In contrast, the Independent Network (IN) model, proposed by Caramazza (1997), and Caramazza and Miozzo (1997) regards gender as an inherent feature of nouns that people will use without awareness. Caramazza et al. (1997) claim that the selection of a noun's grammatical gender is assumed to be an automatic and non-competitive process; this process follows the selection of the lexical form node. Semantic representations can activate word forms directly, without assuming an intervening lemma node.

In order to investigate whether grammatical gender is an important node activated during word production, some researchers focused on how bare nouns and nouns with

determiners are produced. That is, when a grammatical gender affects the bare-noun-production in a similar way as the determiner-noun-phrase production, grammatical gender is used out of awareness during the word production, which confirms the Independent Network. If, however, there is significant difference between the bare-noun-production and the determiner-noun-phrase production, three cognitive layers are activated subsequently, and one-way activation spreading from the lemma to its syntactic frame, confirming the "lemma model".

By using a picture-word interference paradigm, Schriefers and Teruel (2000) claim that three major processes are involved in the production of German noun phrases which include determiner, adjective, and noun: the selection of the noun, the selection of the adjective, and, the selection of the noun's grammatical gender. These three processes appeared at different SOAs (Stimulus Onset Asynchrony). The participants were instructed to name colored line drawings of common objects by means of noun phrases (NPs) like "the red table" as quickly as possible. The distractor words, which the participants were explicitly instructed to try to ignore, were presented by headphone. To each picture, there were seven different distractor conditions[1]. The semantic interference effect for noun distractors (at SOA -150ms) was found to appear much earlier than the effect of the adjective distractors (at SOA +150ms) and the gender interference effect (+150 ms). Schriefers and Teruel claim that the semantic interference effect and the gender interference effect could occur at different SOAs.

[1] (a) Semantically related noun, gender congruent (SEM-CON), (b) semantically related noun, gender incongruent (SEM-INC), (c) semantically unrelated noun, gender congruent (UNR-CON), (d) semantically unrelated noun, gender incongruent (UNR-INC), (e) color adjective (COL-A), (f) unrelated adjective (UNR-A), and (g) no distractor (NONE).

Even without the effect of the adjective distractors (white line drawings were used), the semantic interference effect occurred still earlier than gender interference (the semantic interference effect vanished at SOA +75 ms while a gender interference effect appeared). Results in this research confirm that selecting a target noun does not automatically imply the selection of the corresponding gender feature (Schriefers & Teruel, 2000). The gender feature is selected only when a gender agreement target in the noun's syntactic context was determined. It also shows that the semantic features and grammatical gender are activated in different stages of noun production. It will be discussed that how semantic features and grammatical gender are activated again later.

In studies by Vigliocco et al. (2004), a continuous picture naming paradigm is used to investigate the role of grammatical gender and semantics in German noun production. Participants were instructed to produce bare nouns or nouns with either definite determiners (*die* for feminine, *der* for masculine, or *das* for neutral) or indefinite determiners (*ein* for both masculine and neutral, or *eine* for feminine). It showed that semantic similarity effects occurred when producing either bare German nouns or German noun phrases. That is, intended words to the displayed picture and the intruder (error producing word) were similar in meaning (e.g. saying "Schaf (sheep)" when "Ziege (goat)" is intended); however, target and intruding words tended to share the same gender to a higher degree only when the noun was uttered in a gender-marked context (Vigliocco et al. 2004). This reaction is considered a tendency to preserve the target gender. This preservation did not occur either when producing bare nouns or when producing masculine (or neuter)-indefinite-determiner

noun phrases. Rather, it occurred when producing definite-determiner noun phrases and feminine-indefinite-determiner noun phrases. Vigliocco et al. argue that, in their results, gender-marked frames are highly activated and retrieved, whether or not the corresponding lemma is selected for production. Such a finding is difficult to accommodate in the "lemma model" (Levelt et al., 1999) which allows for one-way activation spreading from the lemma to its syntactic frame only; thus, the gender-marked syntactic frame cannot be retrieved before the lexical selection is completed.

Furthermore, studies on patients who are deficit in retrieving names were reported. For example, Badecker et al. (1995) reported a case with the participant called Dante who is an Italian and sometimes encounters word-finding difficulties because of sequela of meningoencephalitis. He did not have access to any phonological information about a word when unable to name it, despite his ability to distinguish the target's grammatical gender. Since there are some transparent endings that can be gender cues in naming task (e.g. ending /o/ for masculine gender, ending /a/ for feminine gender), but some are not (e.g. ending /e/, /i/, and /u/), it was investigated that lemma rather than phonological forms do play an essential role in producing nouns. Picture-naming tasks and sentence-completion task were involved in this research. On each trial, as Dante was unable to name or produce a word that he nonetheless seemed to know, he was asked to identify the grammatical gender of the target by pointing to one of two cards with the labels "masculine" and "feminine". It showed that Dante was equally good at identifying the gender of regular (ending /o/ or /a/) and exceptional (ending /e/, /i/, and /u/) nouns that he could not name nor otherwise recover

phonological details of. In other words, Dante was able to retrieve word-specific grammatical information even when he was completely unable to provide any indication about the phonological or orthographic form of a lexical item. Thus, the two-stage model of lexical retrieval is proven.

Macoir and Béland (2004) reported another single case. In this case, word production models, the "lemma model" and Independent Network (IN) model, were discussed by testing the responses from a jargonaphasic patient (BA), who is a native speaker of French. Because the difference between "lemma model" (Levelt *et al.*, 1999) and the "Independent network" (IN) model (Caramazza, 1997) is whether or not the gender information is accessed before the sound or written form of the word is retrieved, Macoir et al. tried to find out whether the BA can know the word's gender without knowing the specific word. In French, nouns are assigned to one of the two grammatical genders, masculine or feminine. Determiners (e.g. the indefinite articles "un" and "une" are assigned to masculine and to feminine nouns, respectively) are applied according to the grammatical genders. In this research, the patient's performance in gender identification was expected to differ according to the presentation of a morphological (un/une), for, according to the "lemma model", the syntactic properties of a word is recovered during lemma retrieval, at the first step of lexical retrieval. Indeed, BA could match the noun with un/une on the basis of memorized lexical associations without accessing the gender of the noun at the syntactic level (Macoir & Béland, 2004). The research then explores the access to the syntactic properties of words from the semantic system by means of a gender-decision

task performed on nouns evoked by pictures. Results suggest that the retrieval of the grammatical gender was largely preserved in BA, and did not differ according to the gender-agreement markers. In addition, responses were independent of the correct activation of the corresponding phonological and/or orthographic forms. BA's ability to identify correctly the grammatical genders of words for which she produced spoken errors was good, even when the production was not target-related.

### 1. 3. 2. Double Selection Model

Though there is much evidence supporting the view of competitive activation of the lemma note, recent models are deficient in explaining all phenomena found by the recent research. Results reported by Cubelli et al. (2005) show this deficiency: They report a series of picture-word interference experiments with Italian-speaking participants and find different results for the bare nouns, in which a consistent and robust effect of grammatical gender in the production of bare nouns occurred. In order to investigate whether grammatical gender is activated when people produce bare nouns, and nouns with definite articles, the researchers set up a series of targets that were elements of diverse semantic categories and were presented as line drawings, half of which were masculine, and the other half feminine. Participants were instructed to name the targets as soon as possible when the target was displayed with a distracter. The distracters were either of the same semantic categories and the same grammatical gender, or the same semantic or the same grammatical gender, or neither. In this naming task, semantic category effects were found which strongly supported the

lexical access model. Cubelli et al. also found a main effect of grammatical gender congruity. Reaction time (RTs) were longer when target and distracter words shared the same grammatical gender than when they had different genders, whether or not the stimuli with transparent endings were used , and whether -a or -o were used to denote masculine or feminine, or stimuli with the intransparent ending -e. This result is inconsistent with the hypothesis that syntactic properties are not selected in bare noun production. These results were interpreted as reflecting a competitive lexical selection due to an abstract grammatical gender feature rather than to the phonological or morphological similarity of targets and distracters. The pattern of results show that there was an interference effect on naming times of two independent factors, two semantic relatedness, or two gender congruity. The findings failed to consist completely with either the WEAVER++ model (i.e. mandatory selection of a noun's grammatical gender), or with the IN model (automatic gender activation).

Therefore, Cubelli et al. proposed a Double Selection model (Figure 1-1), which is "consistent with the WEAVER++ model (Levelt et al., 1999 and Roelofs, 1992) in postulating that grammatical information is selected before accessing the word form. However, at variance with WEAVER++, it assumes a direct link between semantic representations and phonology. For a language like Dutch, this model predicts no effects of syntactic selection in bare noun production, thus maintaining the functional architecture of the IN model (Caramazza & Miozzo, 1997)" (Cubelli et al., 2005). Namely, it depends on whether or not the features of the language, gender information, are activated automatically (e.g. Italian and Dutch, mandatory and automatically,

respectively).

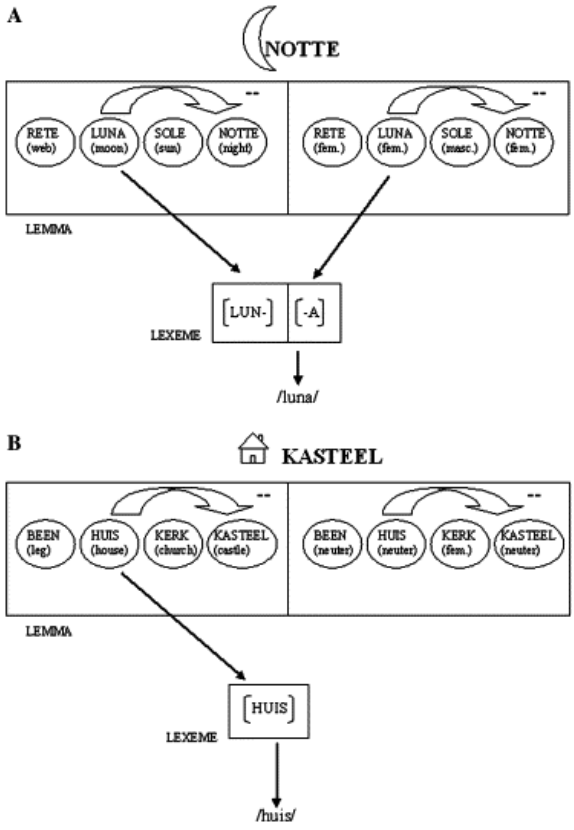



**Figure 1-1 Schematic representation of bare noun production in picture–word interference paradigm.**

(A) Italian; (B) Dutch. Gender interference effect is predicted in Italian but not in Dutch.

### 1. 3. 3. Sex of Referents and Grammatical Genders

As mentioned in the studies above, grammatical gender is regarded as an inherent grammatical feature that is independent of the sex of referents. However, some other research shows that it is not. For example, Vigliocco, Vinson, Paganelli, and Dworzynski (2005) argue that grammatical genders might affect the "similarity judgment" and "picture-naming" with the interference of the sex of referents. In their research, languages with two grammatical genders (Italian), three grammatical genders (German), and no grammatical gender (English) were used; also words with or without sex features (animals or artifacts) were taken into account. In their opinion, grammatical gender might affect cognition by similarity and gender or by sex and gender. From this point of view that similarity and gender influence cognition, gender effects would not depend on the establishing association between the grammatical gender of nouns and the sex of referents. In other words, irrespective of the number of gender categories (e.g. two in Italian, three in German) languages will show a similar grammatical gender effect (similarity and gender hypothesis); By contrast, if cognition is affected by sex and gender (sex and gender hypothesis), effects of grammatical gender could be based on establishing associations between the gender of nouns and sex. In this case, there would be different gender effects across languages (Vigliocco et al., 2005). In the similarity judgment tasks, participants were instructed to choose the two words of the three that were most similar in meaning and to delete the odd word. Italian and German words were compared with English words. This shows that average proportion of same-gender word-pairs are selected significantly higher when

Italian speakers judging the Italian animal words than when English speakers judging the English animal words or when German speakers judging the German animal words. However, when all the words are replaced by pictures, proportion of selecting same-gender animal picture-pairs by Italian speakers was no longer higher than that by English or German speakers. When artifacts were judged, whether the display was a word or picture, no difference was found between the three languages in the proportion of participants who selected same gender pairs. The picture-naming tasks tested whether Italian substitution errors are influenced by grammatical gender, taking into account other factors affecting the likelihood to produce a lexical error. Participants in this study were asked to name the pictures aloud. The results support the "sex and gender hypothesis". That is, grammatical gender effects only occur in the two-gender language when responding to animal words. Vigliocco et al. argue that the form correlation of gender (within the word or in phrasal contexts) plays a role in establishing or strengthening the association between grammatical genders of the words and conceptual features (such as male or female properties) during language development. In contrast to Italian, there is less form correlation of gender in German. The lack of form correlation, combined with the less transparent link between genders of nouns and the sex of human referents, may explain the different results of the two languages.

In summary, there are several essential features of grammatical gender effects. On the facet of lexical access, it might be an inherent property of nouns that can be retrieved at the moment of lexical access for words presented out of context (Bates et

al., 1995). Endings of nouns, which might be a cue of the grammatical gender, can induce gender bias (supported by Italian, French, and German). Such an effect might be a "gender-to-ending direction" or an "ending-to-gender direction" (Bates et al., 1995; Bölte et al., 2004; Colé, Pynte, & Andriamamonjy, 2003). This does not alleviate arguments regarding "conscious awareness". Whether mental awareness of the gender dimension is required during the lexical access is still debatable (Bates et al., 1996; Bentrovato et al., 2003). Such awareness can have interaction with other lexical effects, for instance cohort-effects (Dahan et. al., 2000). Another feature to consider is that grammatical gender is a necessary property of nouns that are masculine or feminine, even neuter in some languages. Normally it bears no relation to the sex of the referent but does affect our mental processing. Nevertheless, according to the "sex and gender hypothesis", grammatical gender effects only occur in the two-gender languages when processing animal names, but does not occur in three-gender language nor when in processing artifacts (Vigliocco et al. 2005).

On the facet of nouns producing processes, two-stage model, "lemma model" (Levelt et al., 1999; Roelofs, 1992) and "Double Selection model" emphasized the essential role played by grammatical gender in producing nouns. Although "Independent Network (IN)" (Caramazza, 1997; Caramazza et al., 1997) argues an automatic activation of grammatical gender, once syntactic context (determiners or adjectives) is provided, the gender frame is needed to activate grammatical gender. That is, grammatical gender is mainly considered to be an independently activated node (Badecker, Miozzo, & Zanuttini, 1995; Levelt, Roelofs & Meyer, 1999; Macoir &

Béland, 2004; Roelofs, 1992; Schriefers & Teruel, 2000), which can be recorded and analyzed separately; moreover, forms of gender (either appropriate determiners or suffixes) are argued to be essential cues during either lexical accessing or noun production, according to which nouns might be grouped in memory tasks. When nouns are learned with the definite article (noun phrase), in nominative for example, feminine noun phrases (NPs) always have the same onset *die*, whereas masculine NPs have the onset *der*, and neuter NPs *das*, which produce more similarities among NPs belonging to the same group.

Therefore, in my opinion, grammatical gender might be an essential mental category that can be used as one of the cues during recognition tasks besides semantic similarity and phonological similarity. However, according to Vigliocco et al. (2005), grammatical gender effects on word production is relative to sex of referents, which cannot be found in three-gender languages (e.g. German). Thus, in this paper, it has been investigated whether grammatical gender can affect human memory. More accurately, it has been discussed if grammatical gender information is used as a memory cue in recognition tasks. An overlap in grammatical gender between materials in the learning stage and the recognition task may result in a higher proportion of false alarms to materials presented only in the recognition task. In addition, my experiment was first run with a two-gender condition (feminine and masculine), and then extended to a three-gender condition (feminine, masculine, and neuter); so that it can be discussed, whether "similarity and gender hypothesis" can be found in recognition tasks. If grammatical gender effect occurs only in the two-gender

condition but not in the three-gender condition, “sex and gender hypothesis” is proved; otherwise, if grammatical gender effect occurs in the three-gender condition either, “similarity and gender hypothesis” is prove.

## 1. 4. Overview of Present Recognition Studies

Thus far, evidence from lexical access and noun production studies has shown robust effects of grammatical gender. Though gender-marked endings are cues of grammatical genders, gender rules of the mother language are indeed playing an essential role in lexical access; in addition, activation of semantic features and the activation of grammatical gender belong to two different stages. However, as mentioned above, recent research still lacks clear support that people are affected by grammatical gender in memory tasks. Since grammatical gender is an individually activated stage that affects lexical access and word production, it could also influence the memory.

The roles of phonological similarity, semantic category, and of semantic association to false memory, are shown by the following discussion. From these studies, a method to explore the influence of grammatical gender on false memory is discussed.

Studies of false memory on semantic lexicon can thus far be assigned two kinds of models: “Feature-based semantic arrangement models” and “Associated semantic arrangement models”. The former model emphasizes the features shared by the learned items and non-presented items. For example, CAT and FOX are near-semantic

neighbors because they share a number of features (e.g., four legs, fur, tail, etc.). Activation of the recognition of item "CAT" can facilitate the features of its near neighbor "FOX" to become activated. By contrast, in the "Associated semantic arrangement models", some special conceptual links exist between PIE and APPLE. Thus, although they do not have obvious feature overlap, they are highly associated and, as so, reside in the same semantic neighborhood. Activation can be spread to one representation, once the other one is activated (Buchanan, Brown, Cabeza, and Maitson, 1999).

### 1. 4. 1. Studies with Associated Semantic Arrangement Models

One of the paradigms, which support the "Associated semantic arrangement models", is called Deese-Roediger-McDermott (DRM) paradigm. This paradigm consists of two tasks: the learning task and the free recall task. In the learning task, people learn a list of words including several groups. For example, *bed*, *rest*, *awake*, *tired*, *dream*, *wake*, *snooze*, *blanket*, *doze*, *slumber*, *snore*, *nap*, *peace*, *yawn*, *drowsy*, etc. are words in the group that associate to the word *sleep*, while *sleep* is the non-present critical item. In the free recall task, the non-present critical items (e.g. *sleep*) will be recalled falsely with about the same proportion as actually present items from the middle of the learned list (e.g. *bed*, *rest*, *awake*, *tired*, etc. as the example list above) (Watson, Balota, & Roediger III, 2003). This effect could not be attenuated or eliminated in free recall tests, even when participants had the chance to correct errors (McDermott, 1996).

McDermott (1996) set up two types of orders to display the learned word list:

blocked and random. For blocked lists, sets of associates were arbitrarily assigned to positions within the list; while for random lists, items were randomly assigned to positions within the list. According to McDermott (1996), the semantically associated blocks of words might encourage relational processing and therefore be more likely to elicit the semantically related words as false alarms. The result, in line with findings by Toglia, Neuschatz, Goodwin, and Lyon (1995a, 1995b, see McDermott, 1996), show that greater false recall occurs after being blocked than after random conditions. Nevertheless, the repetition of the words in the word list did not reduce the error rate of the false recalls.

The following paradigm is extended to phonological similarity between target and distractor words (McDermott, & Watson, 2001; Sommers, & Lewis, 1999; Watson, Balota, & Roediger III, 2003). The study by McDermott (1996) already showed that other items, besides semantically related ones, were falsely recalled by participants, such as phonetically similar words, in free recall. Sommers and Lewis (1999) stated that false memories might be caused by phonological as well as semantic similarity between target items and distractors. In the following research by McDermott et al., (2001), both semantic and phonological effects were explored. McDermott et al. used two types of lists: semantic list and phonological list. For instance, for the critical non-presented item "*Ball*", there was a semantic list including *bounce*, *throw*, *basket*, *bowling*, *golf*, etc., and a phonological list including *doll*, *bile*, *bail*, *balk*, *wall*, *fall*, etc. After learning either semantic or phonological list, participants were instructed to write down recalled words to each list. The results showed that studied items were recalled

more often than semantically related new items (critical non-presented words), while phonologically related new items were recalled with greater proportions than studied items. Proportions of accurate recall increase with increased presentation duration (PD). Proportions of false recall (FR) increase with increased presentation duration.

Watson et al. (2003) extended the results from McDermott et al. (2001): there is a dramatic increase in false recall after simply adding one, two, or three phonologically related words in the learned word lists; while, in contrast, there is a slight increase in false recall with three semantic associates are added. Except for the pure semantic or pure phonological studies, a hybrid list, which includes both semantic and phonological related words, is used. Combined presence of semantic and phonological associates produce over-additive influences on false memory performance. According to Watson et al., these results indicate that there are more false recalls when there is convergence from both conceptual and perceptual (phonological) processing domains; individuals are particularly susceptible to false memories when there is a converging influence of relatively independent dimensions of stimuli (meaning and pronunciation) on non-presented events.

Dodd and Macleod (2004) claim that false memory with DRM paradigm can happen without intentional learning. In their research, an unexpected memory test was applied after the color identification task (unintentional learning phase) or a traditional leaning task (Experiment 1 and 2). The authors found that false alarms were high in both conditions.

### 1. 4. 2. Studies on Feature-Based Semantic Arrangement Models

"Feature-based semantic arrangement models" focus on the relationships between semantic category and recognition tasks (Dewhurst & Anderson, 1999; Dewhurst, 2001; Dewhurst & Farrand, 2004; Roediger & McDermott, 1995; Gardiner, 1988; Tulving, 1985). Recent research has shown that participants in recognition experiments can experience illusory memories of non-studied items if they are semantically related to previously studied items (Dewhurst, 2001). In the earlier research about recognition memory, participants were asked to make "Remember-Know" decisions to each recognized items, so that false memory was detected (Gardiner, 1988; Tulving, 1985). In this "Remember-Know" procedure, participants categorize an item as a "remember" response if they can consciously recollect seeing the item at study, or a "know" response if the item appears to be familiar but they cannot recollect its earlier presentation. High proportion of false alarms was categorized as R responses (Roediger & McDermott, 1995); however, K responses were found enhanced by semantic processing (Gardiner, 1988). Researchers started to argue that remember and know responses reflect qualitatively distinct aspects of recognition memory (see Dewhurst, 2001, for review). Remember responses are sensitive to manipulations that engage conceptual processes, whereas know responses are sensitive primarily to manipulations of perceptual properties of stimuli (Gardiner, 1988). However, Conway, Gardiner, Perfect, Anderson, and Cohen (1997) argue that know responses might be sensitive to conceptual as well as perceptual manipulations by asking participants to categorize their answers to multiple-choice questions as

"remember," "just know," "familiar," or "guess". By using similar multiple-choice questions as "remember", "know", or "guess[2]", Dewhurst and Anderson (1999) argue that results in their research could not be accounted for by the view that know responses reflect perceptual processes. Dewhurst et al. (1999) set up two kinds of repetition conditions: the exact repetition condition and the category repetition condition. Thirty-six different semantic categories are chosen (e.g., "musical instruments", "parts of body", etc.). In the exact repetition condition, one target item from each semantic category is chosen and each target item is presented either one, four, or eight times in succession; in the category repetition condition, one, four or eight target items in each of 36 categories are presented. It shows that exact repetition influences recognition performance by enhancing the recollection component of recognition memory, as measured by remember responses. By contrast, category repetition influences recognition memory by enhancing the familiarity of both studied and non-studied category members, resulting in increases in both correct and false positive know responses. Dewhurst et al. interprets the increasing know responses by category repetition as spreading activation within a semantic network. Category repetition leads to the false recollection of non-studied category members (Dewhurst & Anderson 1999).

Furthermore, semantic features and semantic associations are found to make different contributions to false memory (Buchanan et al., 1999). Participants received either semantic feature-based items (category list) or semantic associated items.

[2] There may be other items that participants neither recollect nor recognize on the basis of familiarity, but which they cannot definitely reject. Participants have the option of making a guess response to these items if they wish.

Buchanan et al. found that "participants in the association study condition were much more likely to falsely recognize the non-presented items than those that were subjects in the category list." Aside from the influence to false memory from feature or category links, the authors claim a greater spread of activation to non-presented but associatively related items.

In short, false memory can be raised when non-studied items share semantic associations or semantic and/or phonological association features with the studied items. Remember responses are sensitive to semantic features, while know responses are sensitive to semantic or phonological features when different research conditions are involved (R-K choice or multiple-choice).

## 2. Intention of My Research

In this paper, I am interested in whether grammatical gender influences false memory beyond phonological and semantic similarity. As discussed, previous research shows strong effects of grammatical gender on both lexical access and word production, either with awareness or competitive selection, through automatic activation. Grammatical gender is considered to affect our cognitive process deeply and independently. However, recent research lacks clear evidence that people are affected by grammatical gender in memory tasks. Since phonological features and semantic features have strong effects on recognition tasks, false memory can be raised when non-studied items (lures) share semantic associations or semantic and/or phonological features with the studied items (targets); it is also expected that grammatical gender

plays an essential role in recognition as it is one more feature that can be used as a cue to reject or accept a non-studied word. Namely, false memory will be raised by grammatical gender cues once the lure and the target share the same grammatical gender information. Influences of grammatical gender markings, which were found important on lexical access and word production, are also relevant in my research. I am interested in whether rules of grammatical gender can be activated with or without gender markings in false memory tasks.

As mentioned above, recent studies on false memory show that the more features targets and lures share, the higher possibility for false memory to occur. Therefore, in my study, the experimental paradigm was based on Dewhurst et al. (1999). Targets and lures belonged to the same semantic category and may or may not share the same grammatical gender feature. Aside from the semantic category effect, it is assumed that influences on false memory from grammatical gender exists in the recognition task if differences between the results of correct rejection to the lures shared and did not share the same grammatical gender feature with targets.

Grammatical gender, in my research, is expected to affect recognition in a similar way as found for semantic similarity by Dewhurst et al. That is, when all the lures belong to the same semantic category as targets, more false alarms will occur to the lures which share the same grammatical gender information as targets (effects from grammatical gender similarity); by contrast, lures that do not bear the same grammatical gender, as targets are affected only by semantic similarity, are expected to be more easily rejected correctly. Similar effects to semantic similarity on targets found

by Dewhurst et al. (1999) across different grammatical genders are expected.

## 3. Experiment 1

In this Experiment, effects of grammatical gender in recognition memory in German are investigated. Grammatical gender was expected to influence recognition memory, as lure words from the same gender category as target words share more features with a target word and should therefore be more difficult to reject as a lure word. In addition to the well established effect of semantic category (Dewhurst, et al., 1999, 2001, and 2004; Dodd et al., 2004; Gardiner, 1988; Roediger et al., 1995; McDermott, 1996; McDermott, et al., 2001; Tulving, 1985; Sommers, et al., 1999; Watson, et al., 2003), an effect of grammatical gender on recognition memory was predicted.

In the main study, the influence of relations between the grammatical gender of targets and lures was investigated. Feminine nouns served as targets or lures, masculine nouns served as lures unrelated in terms of grammatical gender.

Recent research shows that word recognition tasks are influenced by different factors, for instance semantic category effects. To fulfill my research, I arranged targets and lures in order to control potential effects of semantic similarity. For example, within the semantic category of birds, a noun such as *sparrow* is more closely related to *pigeon* than *penguin*. When *sparrow* is presented as the target, the word *penguin* might be more easily rejected as a lure word in a recognition task than *pigeon*. A pretest was set up before the main study, for controlling representative and typical similarity of words within each semantic category, so that recognition tasks were

influenced by neither semantically representative nor typical effects of words.

## 3. 1. Pretest of Experiment 1

As the materials used in the main study were selected mainly from the word list used by Dewhurst et al. (1999), the pretest ensured semantic similarity in German and comparable semantic similarity of both feminine lures and masculine lures with targets.

### 3. 1. 1. Method

**Participants:** 56 native German speakers took part in this pretest via internet voluntarily.

**Materials:** Thirty-six semantic categories of three members each were selected from the materials used by Dewhurst and Anderson (1999), including parts of the body, birds, etc., or were compiled from dictionaries. In each category, two feminine words and one masculine word were selected, for example in category *Units of time*, two feminine words *Woche* (week), and *Stunde* (hour), and one masculine word *Tag* (day). In the word list from Dewhurst and Anderson (1999), there were some categories where nouns belonged to the same grammatical gender, for example names of city etc.; these categories were replaced with categories whose items vary in grammatical gender, such as emotions (Appendix A). All materials can be found in *Das neue Deutsch-Chinesische Wörterbuch* (The new German-Chinese dictionary, first edition, 2000), and were double checked by three German native speakers so that words with ambiguous gender in spoken language were avoided. All words had similar frequencies

as checked in the corpus via internet. Class of frequencies varied between 6 and 16 (the word "*der*" is approximate 2^6 till 2^16 times more frequent than the selected item); all items had comparable class of frequencies within the same categories (Wortschatz Universität Leipzig)[3]. All words of similar frequencies were combined with each other resulting in 3 word pairs within each category (Stunde-Woche/hour-week, Stunde-Tag/hour day, Woche-Tag/week-day). Therefore, a word list with 108 word pairs was achieved.

### 3. 1. 2. Procedure

Words from each semantic category were presented as word pairs, for example, hour-week (Stunde [fem.1]-Woche [fem.2]), hour-day (Stunde [fem.1]-Tag [masc.]), week-day (Woche [fem.2]-Tag [masc.]) etc., resulting in a word list with 108 word pairs; two feminine words in each semantic category were arranged into [fem.1] or [fem.2] at random. In order to limit the task to an acceptable length, each participant rated 27 word pairs (a quarter of 108 word pairs). All word pairs were followed by a 7-point rating scale (1 = no relationship between the two words, 7 = close relationship). The pretest was run as an online study.

[3] *http://wortschatz.uni-leipzig.de/*

### 3. 1. 3. Results

All ratings were analyzed by a one way ANOVA with the factor of word-pair groups (fem.1-fem.2, fem.1-masc., and fem.2-masc.) at an alpha level of 0.05. The results showed that similarity between word pairs was not influenced by grammatical gender across different semantic categories of word pairs, $F_{(2, 3022)} = 1.80$, $p = 0.17$. Mean scale of judging directions of word-pair relationships between "[fem.1] and [masc.]" ($M$ =5.21, $SD$ =1.56) was not only close to that of that of "[fem.1] and [fem.2]" ($M$ =5.30, $SD$ =1.40), but also close to that of "[fem.2] and [masc.]" ($M$= 5.28, $SD$= 1.48). The paired sample t-test showed no significant difference among different group pairs: judgments of word-pairs "[fem.1] and [masc.]" were not significantly lower than that of "[fem.1] and [fem.2]", $t_{(1511)} = 1.85$, $p = 0.06$; meanwhile, significant differences were not found between judgments of word-pairs "[fem.2] and [masc.]" and "[fem.1] and [fem.2]", $t_{(1511)} = 0.26$, $p = 0.80$, or between judgments of word-pairs "[fem.1] and [masc.]" and word-pairs "[fem.2] and [masc.]" $t_{(1511)} = 1.47$, $p = 0.14$. All words selected can therefore function as targets or lures in the main study. In the main study, feminine words from the fem.1-list and masculine words were used as lures and feminine words from the fem.2-list were used as targets

## 3. 2. Main Study of Experiment 1

### 3. 2. 1. Method

**Participants:** Forty-two[4] native German speakers from Heidelberg University (11 male, 31 female, mean age is 23.47, *SD* = 3.87) were involved in the main Experiment.

**Materials:**

In this experiment, feminine words served as targets or lures, masculine words served as lures. There were two kinds of semantically related lures: (1) feminine related lures (gender-related lures) shared the same semantic category and the same grammatical gender with the targets; (2) masculine related lures (gender-unrelated lures) shared the same semantic category but had different grammatical gender than the targets. Both of these two kinds of lures occurred in the recognition phase only. Thirty-six filler items were used as semantically unrelated lures. Half of the filler items were feminine and the other half were masculine. Participants either received the experimental stimuli in a massed repetition or a spaced repetition type. In the massed repetition type each item was presented either once, four, or eight times in succession. In the spaced repetition type, repetitions were separated by at least three intervening items. The recognition test consisted of 36 feminine targets, 36 feminine lures, 36 masculine lures, and 36 filler items.

Recent research shows a number-of-repetition effect in word recognition tasks (Dewhurst & Anderson, 1999; Dewhurst, 2001; Dewhurst & Farrand, 2004): the more time the word, or the words within the same semantic category, is repeated, the more

[4] The sample size of 42 yields a statistical power of .9962, given an effect size of .25 and an alpha-level of .05 (analysis based on G*Power; see Faul, Erdfelder, Lang & Buchner, 2007).

false alarms to non-studied category members are found. Therefore, it is assumed that repetition might produce an increase in false positive responses to the gender-related lures, which also share the same gender feature as the studied members, as well as inhibit the false alarms to the gender-unrelated lures. Thus, number of repetitions was varied within-subjects, with each participant seeing 12 items once, 12 items four times and 12 items eight times. Repetition of items was rotated through three study lists so that each item appeared at each level of repetition for an equal number of participants. Test items were presented in a single random order on four sheets of A4 paper, each consisting of two columns of 18 items, with the letters R (*remember)*, K (*know)*, G (*guess)*, and No printed to the right of each item. The dependent measures were the number of hits and false alarms designated as R, K, and G responses.

### 3. 2. 2. Procedure

Study items were presented on Apple Macintosh computers. Each word remained on the screen for one second and was replaced after an interval of one second by the next item. Participants were instructed to read the words silently as they appeared on the screen and to bear in mind that they would later be given some form of memory test, the precise nature of which was not specified.

**Figure 3-1 Rush hour Game**

Presentation of the 156 trials took approximately ten minutes. Participants were then provided with a distractor task for 15 minutes[5] (Figure 3-1).

Following the distractor task, participants were given a recognition test. They were asked to identify the words they had learned: If they believe that an item did not appear in the earlier phase, they should answer "*No*". If they believe that an item did appear in the earlier task, they should circle either "*Remember*", "*Know*", or "*Guess*", which refers to the nature of their conscious experience as they recognized the item (Dewhurst et al, 1999).

Participants were instructed to give an R response only if they can recollect the presentation of the item. All participants were asked to work through the response

[5] Here a problem solving task called *Rush Hour* was used, which is a puzzle task with a target car in the parking lot. Participants were instructed to move the cars parking around the target car so that the target car can be driven out of the parking lot.

sheets once, without returning to previous sheets, beginning at the top of the left-hand column on each page.

### 3. 2. 3. Results

Response proportions of targets, lures, and fillers are displayed in Table 3-1, Table 3-2, and Table 3-3. To the lures and fillers, all the *Remember*, *Know*, and *Guess* responses were false alarms, while "No" responses were correct rejections; by contrast, to targets, all the *Remember*, *Know*, and *Guess* responses were hits, while "No" responses were misses.

**Table 3 - 1 Response Proportions (*p*) and Standard Deviations as a Function of Number of Repetitions and Repetition Type for Semantic lures (Correct Rejections only)**

| | | Repeated 1 time* | | Repeated 4 times | | Repeated 8 times | |
|---|---|---|---|---|---|---|---|
| | | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | F | 0.90 | 0.13 | 0.89 | 0.12 | 0.89 | 0.09 |
| | M | 0.95 *u* | 0.07 | 0.94 *u* | 0.09 | 0.96 *u* | 0.06 |
| Massed repetition | F | 0.87 | 0.13 | 0.83 | 0.16 | 0.89 | 0.07 |
| | M | 0.93 *u* | 0.08 | 0.90 *u* | 0.06 | 0.92 *u* | 0.05 |

Note: F: feminine lures, M: masculine lures.

Since the correct rejections were mainly focused in this Experiment, only data of correct rejections were showed in this table.

**Repetition* here refers to the number of presentations of the target items from the same semantic category.

Data with *u* were proportions of correct rejections to the gender-unrelated lures.

**Table 3 - 2 Response Proportions (*p*) and Standard Deviations as a Function of Number of repetitions and Response Type for Targets**

| | Targets (feminine) | | | | | |
|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | |
| Remember | 0.67 | 0.10 | 0.88 | 0.11 | 0.92 | 0.14 |
| Know | 0.08 | 0.11 | 0.04 | 0.06 | 0.02 | 0.05 |
| Guess | 0.03 | 0.07 | 0.04 | 0.07 | 0.03 | 0.07 |
| No | 0.22 | 0.14 | 0.04 | 0.06 | 0.03 | 0.07 |
| Massed repetition | | | | | | |
| Remember | 0.71 | 0.13 | 0.85 | 0.10 | 0.82 | 0.13 |
| Know | 0.11 | 0.14 | 0.05 | 0.05 | 0.08 | 0.10 |
| Guess | 0.08 | 0.08 | 0.04 | 0.06 | 0.09 | 0.09 |
| No | 0.10 | 0.12 | 0.06 | 0.07 | 0.01 | 0.09 |

**Table 3 - 3 Response Proportions (*p*) and Standard Deviations as a Function of Number of repetitions and Response Type for Filler items**

| | Filler items* | | | | | |
|---|---|---|---|---|---|---|
| | Feminine fillers | | Masculine fillers | | Across genders | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | |
| Remember | .05 | .07 | .08 | .07 | .06 | .06 |
| Know | .02 | .06 | .02 | .04 | .02 | .05 |
| Guess | .13 | .11 | .06 | .08 | .10 | .09 |
| No | .80 | .09 | .84 | .07 | .82 | .09 |
| Massed repetition | | | | | | |
| Remember | .06 | .08 | .05 | .06 | .05 | .06 |
| Know | .04 | .07 | .04 | .08 | .04 | .07 |
| Guess | .10 | .09 | .10 | .10 | .10 | .09 |
| No | .80 | .06 | .81 | .08 | .81 | .07 |

The analysis focused on the influence of grammatical gender on detecting lure words, that is, on the response proportions of feminine lures (gender-related lures) and masculine lures (gender-unrelated lures).

An alpha level of .05 was used in all statistical analyses.

### 3. 2. 3. 1. Correct Rejections to Lures

The experimental factors of correct rejections resulted in a 2 (gender conditions: gender-related vs. gender-unrelated lures) x 2 (types of repetition: massed repetition vs. spaced repetition) × 3 (number of repetitions: targets repeated one, four, or eight times) mix ANOVA with number of repetitions and gender conditions being varied within participants and types of repetition being varied between participants (Table 3-4). False alarms occurred rarely (most of the proportions of false alarms were less than 0.11), and were analyzed further.

**Table 3 - 4 Results from the mixed ANOVAS**

**Test of Within-Subjects Effects for Semantic lures (Correct Rejections only)**

| | Sum of Squares | *df* | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Lures | 0.131 | 1 | 0.131 | 15.086 | 0.000* |
| Lures*Type | 0.010 | 1 | 0.010 | 1.183 | 0.283 |
| Error (Lures) | 0.348 | 40 | 0.009 | | |
| Number of repetitions | 0.026 | 2 | 0.013 | 1.506 | 0.228 |
| Times*Type | 0.005 | 2 | 0.003 | 0.316 | 0.730 |
| Error (Times) | 0.680 | 80 | 0.009 | | |
| Lures*Number Number of repetitions | 0.002 | 2 | 0.001 | 0.168 | 0.845 |
| Lures*Number Number of repetitions*Type | 0.007 | 2 | 0.003 | 0.591 | 0.546 |
| Error (Lures*Times) | 0.455 | 80 | 0.007 | | |

Note: *: $p \leq 0.05$.

The analysis revealed a reliable main effect of grammatical gender. Feminine

lures (gender-related lures) were more often falsely accepted than masculine lures (gender-unrelated lures), $F$ (1, 40) = 15.09, $MSe$ = 0.13. This effect occurred in both spaced and massed repetition. The difference between proportions of correct rejections to spaced repetition and massed repetition was marginally significant at the alpha level of 0.05, $F$ (1, 40) = 3.14, $MSe$ = 0.09, $p$ = 0.08. Moreover, the proportion of correct rejections did not change with the variance of number of repetitions, F (2, 80) = 1.51, $MSe$ = 0.01. There were no reliable interactions between experimental factors, $F$'s ≤ 1.18.

Analyzing response proportions for filler items revealed no influence of their grammatical gender, $F$ < 1.

### 3. 2. 3. 2. Hits to Targets

The experimental factors of Hits resulted in a 2 (types of repetition: massed repetition vs. spaced repetition) × 3 (number of repetitions: targets repeated one, four, or eight times) mix ANOVA with number of repetitions being varied within participants and types of repetition being varied between participants. Misses of the targets were rare (most of the proportions of miss were less than 0.10), and were not further considered.

Correct *Remember* responses increased significantly with number of repetitions, $F$ (2, 80) = 42.05, $MS$e = 0.41. One-time repetition resulted in less hits than both four- and eight-time repetition. The T-test showed significant difference both between one-time repetition and four-time repetition, t (41) = 7.16, $p$ = 0.00, and between one-time repetition and eight-time repetition, t (41) = 6.34, $p$ = 0.00; by contrast, differences

between four-time and eight-time repetition were not found, t (41) = 0.06, $p$ = 0.96. In addition, repetition type and number interacted significantly, $F$ (2, 80) = 6.48, $MS$e = 0.06. The number-of-repetition effect was more pronounced in the spaced repetition than in the massed repetition: items presented eight times were associated with more *Remember* responses in the spaced repetition type than in the massed repetition type, $F$ (1, 40) = 7.94, $MS$e = 0.14; by contrast, there was no effect of massed versus spaced repetition when targets were repeated four times or displayed only once, $F$'s ≤ 1.22. However, across number of repetitions, the main effect of repetition type was not significant, $F$ (1, 40) = 1.50, $MS$e = 0.03.

In contrast to *Remember* responses, *Know* responses were only influenced by number of repetitions, $F$ (2, 80) = 5.96, $MS$e = 0.03. Further t-test showed significantly higher proportions of hits to targets presented once than to targets either repeated four times, t (41) = 2.66, $p$ = 0.01, or repeated eight times, t (41) = 3.01, $p$ = 0.01; the difference between targets repeated four times and eight times was not significant, t (41) = 0.35, $p$ = 0.72. Analyzing correct *Guess* responses revealed neither effects of massed versus spaced repetition, $F$ (1, 40) = 1.18, $MS$e = 0.01, nor effects of number of repetitions, $F$ (2, 80) = 1.28, $MS$e = 0.01. There was no interaction between number of repetitions and repetition type, neither with correct *Know* responses, $F$ (2, 80) = 1.17, $MSe$=0.01, nor with correct *Guess* responses, $F$ (2, 80) =1.49, $MSe$ = 0.01.

### 3. 2. 4. Discussion

Compared with results reported by Dewhurst et al. (1999), which were achieved by a similar experimental design, higher proportions of false alarms to lures (both semantic related lures and semantic unrelated lures) were found in this study with a two-gender experimental condition. Such a result indicated two possible reasons: (1) there were more recognition items in this study than in previous studies, which might have enhanced the task's difficulty; (2) grammatical gender did raise the risk of false alarms.

The second assumption was supported by results found for lures. Recognition memory for nouns was influenced by grammatical gender effects, as well as effects from semantic category. In this study, both feminine lures and masculine lures were semantically related to learned words. In addition, feminine lures shared one more character with learned words, grammatical gender, while masculine lures did not. Higher proportions of false alarms to feminine lures indicated the difficulty on rejecting the lures that shared more characters with learned words. That is, grammatical gender information may be used as a memory cue in the recognition task, as well as semantic category information. Lures were more difficult to reject, once they shared not only the same semantic category but also grammatical gender information with the learned words. In other words, grammatical gender cues facilitated rejecting gender-unrelated lures (e.g. masculine lures in this experiment).

As mentioned before, since semantic category effects influenced both gender-related lures and gender-unrelated lures, grammatical gender was the only influence

on false memory to lures. However, analyzing filler items that also shared the feature of grammatical gender with targets, which were not semantically related, showed no effect on grammatical gender. Grammatical gender's influence was limited to materials that had mentally been organized by semantic category.

The results of hits for targets are in line with the study by Dewhurst and Anderson (1999). By using German nouns, the present experiment showed that semantic category enhances recognition by promoting the conscious recollection of individual instances, since the incidence of *Remember* responses increased with the number of times an item was presented.

In short, aside from the semantic category effect found in recent research, grammatical gender effects were also found in Experiment 1 with masculine and feminine nouns in German. However, as German is a three-gender language, in which there are not only masculine nouns and feminine nouns, but also neuter nouns, in the following studies grammatical gender effects with three gender conditions were explored.

# 4. Experiment 2

In this Experiment, two issues were addressed: (1) grammatical gender effects in a three-gender condition were explored; (2) and equal grammatical gender effects in three-gender condition in recognition memory in German were investigated. In Experiment 1, only two genders were used: feminine and masculine. As a result, two problems were unresolved: (1) in the learning stage, participants remembered only

one gender, which can help participants reject a different gender easily. In other words, a memory task in a two-gender condition somewhat enhanced the two-gender effect in German, but German is a three-gender language. How grammatical gender influences the false memory in a three-gender condition should be further investigated. (2) In the research done by Colé et al. (2003), words ending with "*e*" were less likely to be seen as feminine by participants, though masculine nouns were displayed nearly as frequently as feminine nouns. In Experiment 1, only feminine words were used as targets. Therefore, it was still an open question whether grammatical gender effects exist when the targets are masculine or neuter instead of feminine.

In Experiment 2, the main hypothesis was similar to the hypothesis in Experiment 1: aside from the semantic category effects, people are more likely to reject a gender-unrelated lure than a gender-related lure in the recognition task.

## 4. 1. Pretest of Experiment 2

In this Experiment, within each semantic category, two nouns were used as targets, either feminine and masculine, feminine and neuter, or masculine and neuter. Three nouns were used as semantically related lures, including feminine, masculine and neuter nouns. Before the main study, a pretest was set up to control words' representative and typical similarity within each semantic category, as was done in Experiment 1. This was done so that recognition tasks were influenced by neither semantically representative nor typical effects of words.

### 4. 1. 1. Method

**Participants:** Forty-five Native German speakers participated in the internet test voluntarily.

**Materials:** Eighteen semantic categories (e.g., parts of the body, flowers, etc.), each with six members, were either selected from the materials used by Dewhurst and Anderson (1999) or were created especially for the study. In each category, for example in category *four-footed animals*, two feminine words *Katze* (cat), *Ziege* (goat), two masculine words *Hund* (dog), *Elefant* (elephant), and two neuter words *Lamm* (lamb), *Pferd* (horse) were selected. All category members were nouns with equal frequency. In the word list from Dewhurst and Anderson (1999), there were some categories which carried no gender information, for example names of cities; these categories were replaced by other categories in this study, such as emotions (Details of materials are in Appendix B - 1). All materials can be found in *Das neue Deutsch-Chinesische Wörterbuch* (The New German-Chinese Dictionary, 2000, republished in 2001), and were double-checked by three native German speakers so that words with ambiguous genders, when spoken, could be avoided; class of frequencies varied between 6 and 16 (the word "*der*" is approximate 2^6 to 2^16 times more frequent than the selected item); all items had a comparable class of frequencies within the same categories. All frequencies were checked in the corpus via internet (Wortschatz Universität Leipzig).

### 4. 1. 2. Procedure

Words from each semantic category were presented as word pairs, for example, cat-goat (Katze [fem.1]-Ziege [fem.2]), dog-elephant (Hund [masc.1]-Elefant [masc.2]), lamb-horse (Lamm [neu.1]-Pferd [neu.2]), cat-dog (Katze [fem.1]-Hund [masc.1]), dog-lamb (Hund [masc.1]-Lamm [neu.1]), cat-lamb (Katze [fem.1]- Lamm [neu.1]), goat-elephant (Ziege [fem.2]-Elefant [masc.2]), etc., resulting in a word list with a total of 270 word pairs. In order to limit the task to an acceptable length, each participant rated 90 word pairs. All word pairs were followed by a 7-point rating scale (1 = no relationship between the two words, 7 = close relationship). The pretest was run as an online study.

### 4. 1. 3. Results

The results showed that word pairs have similar relationships (Table 4-1).

Repeat Measure test showed no significant difference among these groups, $F_{(14, 3766)} = 1.50, p = 0.11$. All words that were selected can therefore function as targets or lures in the main study. Targets and lures were selected from this list randomly.

**Table 4-1 Results of the Pretest**

| | N | Mean | Std. Error |
|---|---|---|---|
| 1 | 270 | 4.3889 | 9.166E-02 |
| 2 | 270 | 4.5667 | 9.360 E-02 |
| 3 | 270 | 4.5333 | 9.713 E-02 |
| 4 | 270 | 4.6222 | 8.745 E-02 |
| 5 | 270 | 4.6333 | 8.353 E-02 |
| 6 | 270 | 4.6333 | 8.303 E-02 |
| 7 | 270 | 4.6667 | 7.659 E-02 |
| 8 | 270 | 4.6222 | 7.743 E-02 |
| 9 | 270 | 4.6556 | 7.449 E-02 |
| 10 | 270 | 4.6111 | 8.215 E-02 |
| 11 | 270 | 4.5778 | 8.379 E-02 |
| 12 | 270 | 4.6000 | 8.017 E-02 |
| 13 | 270 | 4.5111 | 8.870 E-02 |
| 14 | 270 | 4.4111 | 9.175 E-02 |
| 15 | 270 | 4.3556 | 8.966 E-02 |

## 4. 2. Main Study

### 4. 2. 1. Method

**Participants:** Ninety[6] native German speakers from Heidelberg University (33 male, 57 female, mean age 22.98, *SD* = 3.22) participated in the main study.

**Materials:**

Two words of differing gender (either feminine-masculine, feminine-neuter, or masculine-neuter) were selected from each category as the targets, which were

[6] The sample size of 90 yields a statistical power of .9999, given an effect size of .25 and an alpha-level of .05 (analysis based on G*Power; see Faul et al., 2007).

presented in the learning phase. These were then used as the targets in the subsequent recognition test. After the target words were selected, four words remained in each semantic category; three of these (one feminine, one masculine, and one neutral) were selected to be the semantically related lures. Similarly to Experiment 1, there were two kinds of semantically related lures: gender-related lures and gender-unrelated lures. In addition, eighteen filler items, which did not share the semantic category with the targets, were set up as unrelated lures. All lures, including gender-related lures, gender-unrelated lures, and filler items, were only displayed in the recognition stage.

In addition, two repetition types were arranged: massed repetition and spaced repetition. Participants in the massed repetition type studied the thirty-six target words, with each item presented either one, four, or eight times in succession. In the spaced repetition type, repetitions were separated by at least three intervening items. The recognition test consisted of the thirty-six targets, thirty-six gender-related lures, eighteen gender-unrelated lures, and eighteen filler items.

**Procedure**: The procedure was the same as the one in Experiment 1. Three stages were involved. An approximate ten-minute learning stage was followed by a fifteen-minute RushHour game, which was used as a distractor task. After the distractor task, the participants were given a recognition test. They were asked to identify the words they had learned: If they believed that an item did not appear in the earlier phase, they were to answer "*No*". If they believed that an item did appear in the earlier task, they were to answer either "*Remember*", "*Know*", or "*Guess*", which referred to the nature

of their conscious experience as they recognized the item (Dewhurst et al, 1999).

In the learning stage, each participant received either feminine and masculine nouns ("fem.-masc." targets), masculine and neuter nouns ("masc.-neu." targets), or feminine and neuter nouns ("fem.-neu." targets). Correspondingly, materials used in the recognition task were called "fem.-masc." lures, "masc.-neu." lures, or "fem.-neu." lures according to the targets presented in the learning stage, although lures received by each participant were the same.

### 4. 2. 2. Results

The analysis primarily focused on the correct rejection of lures (both gender-related and gender-unrelated) and the correct identification ("Remember") of the target words. All "Remember," "Know," and "Guess" responses to the lures are false alarms, while "*No*" responses to the lures are correct responses. Therefore, Table 4-2 and 4-3 show the mean proportions of Remember responses to targets, and "No" responses to lures as a function of number and type of repetitions (complete data are in Appendix B - 3).

**Table 4-2 Response Proportions (*p*) and Standard Deviations as a Function of Number of repetitions and Repetition Type for Targets (Answer "remember" only)**

| | | "fem.-masc." group targets | | | "masc.-neu." group targets | | | "fem.-neu." group targets | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| Spaced repetition | F | 0.30 | 0.74 | 0.82 | / | / | / | 0.41 | 0.74 | 0.87 |
| | M | 0.30 | 0.67 | 0.76 | 0.43 | 0.70 | 0.80 | / | / | / |
| | N | / | / | / | 0.41 | 0.78 | 0.74 | 0.39 | 0.67 | 0.72 |
| Massed repetition | F | 0.38 | 0.54 | 0.46 | / | / | / | 0.36 | 0.42 | 0.63 |
| | M | 0.29 | 0.40 | 0.56 | 0.63 | 0.50 | 0.73 | / | / | / |
| | N | / | / | / | 0.65 | 0.66 | 0.69 | 0.46 | 0.71 | 0.73 |

Note: F: feminine targets, M: masculine targets, N: neuter targets.

**Table 4-3 Response Proportions (*p*) and Standard Deviations as a Function of Number of repetitions and Repetition Type for Semantic lures (Correct Rejections only)**

| | | "fem.-masc." group lures | | | "masc.-neu." group lures | | | "fem.-neu." group lures | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| Spaced repetition | F | 0.95 | 0.87 | 0.90 | 0.87*u* | 0.86*u* | 0.77*u* | 0.90 | 0.85 | 0.77 |
| | M | 0.92 | 0.98 | 0.85 | 0.74 | 0.78 | 0.63 | 0.91*u* | 0.93*u* | 0.84*u* |
| | N | 0.93*u* | 0.85*u* | 0.93*u* | 0.79 | 0.70 | 0.75 | 0.88 | 0.76 | 0.79 |
| Massed repetition | F | 0.83 | 0.83 | 0.78 | 0.77*u* | 0.82*u* | 0.77*u* | 0.86 | 0.62 | 0.48 |
| | M | 0.84 | 0.82 | 0.74 | 0.76 | 0.74 | 0.71 | 0.58*u* | 0.31*u* | 0.84*u* |
| | N | 0.74*u* | 0.73*u* | 0.87*u* | 0.68 | 0.78 | 0.84 | 0.71 | 0.55 | 0.54 |

Note: F: feminine lures, M: masculine lures, N: neuter lures.

Data with *u* are proportions of correct answers from gender-unrelated lures.

All of the possible responses to the lures were analyzed (more details are in Appendix B - 4). An alpha level of .05 was used in all statistical analyses.

### 4. 2 .2. 1. Correct Rejections of the Lures

Statistical analyses consisted of separate 3 (gender of lures was feminine, masculine or neuter) × 3 (gender combination group was feminine and masculine, masculine and neuter, or feminine and neuter) × 2 (Repetition type: massed repetition or spaced repetition) × 3 (Number of repetitions: target word was repeated one, four, or eight times) mixed analyses of variance (ANOVAs).

#### 4. 2. 2. 1. 1. Analysis to Gender-Related Lures and Gender-Unrelated Lures

Interaction between "gender of lures" and "gender combination groups", which indicated difference between proportions of correct rejection of gender-related lures and of gender-unrelated lures, was found, $F(2, 168) = 3.17$, $MSe = 0.06$ (Table 4-4).

**Table 4-4 Results from the mixed ANOVAS**

**Test of Within-Subjects Effects for Semantic lures (Correct Rejections only)**

| | Sum of Squares | *df* | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Lures | 0.179 | 2 | 0.090 | 4.95 | 0.008* |
| Lures*Type | 0.052 | 2 | 0.030 | 1.436 | 0.241 |
| Lures*Group | 0.229 | 4 | 0.060 | 3.171 | 0.015* |
| Lures*Type*Group | 0.243 | 4 | 0.060 | 3.353 | 0.011* |
| Error (Lures) | 3.000 | 168 | 0.020 | | |
| Times | 0.120 | 2 | 0.211 | 11.811 | 0.000* |
| Times*Type | 0.266 | 2 | 0.133 | 7.332 | 0.001* |
| Times*Group | 0.512 | 4 | 0.128 | 7.065 | 0.000* |
| Times*Type*Group | 0.364 | 4 | 0.01 | 5.020 | 0.001* |
| Error (Times) | 0.120 | 168 | 0.020 | | |
| Lures * Times | 0.454 | 4 | 0.144 | 5.95 | 0.000* |
| Lures*Times*Type | 0.715 | 4 | 0.179 | 8.796 | 0.000* |
| Lures*Times*Group | 1.300 | 8 | 0.162 | 7.996 | 0.000* |
| Lures*Times*Type*Group | 0.796 | 8 | 0.100 | 4.898 | 0.000* |
| Error (Lures*Times) | 6.828 | 366 | 0.020 | | |

Note: *: $p \leq 0.05$.

"Times" was the number of repetitions of targets, including 1, 4 and 8 times;

"Type" was the repetition type of targets, including spaced type or massed type;

"Groups" was the genders of the displayed target words, including "feminine and masculine", "feminine and neuter", or "masculine and neuter";

"Lures" was the gender of lures, including feminine, masculine and neuter.

In addition, a main effect was found for "type of repetition", $F(1, 84) = 13.59$, $MSe = 2.65$, and interactions among "gender of lures", "type of repetition" and "gender combination groups" $F(4, 168) = 3.35$, $MSe = 0.06$ (Table 4-4). This indicates

that the recognition task was influenced by repetition types (spaced and massed) differently. Therefore, a further analysis was done with separating the whole data into two groups according to different repetition types (Table 4-5, Table 4-6). It revealed no effects for gender-related versus gender-unrelated (grammatical gender effects) with the data submitted by participants who learned targets displayed massedly, $F < 1$; by contrast, such effects was found when the targets were displayed spacedly, $F(4, 84) = 7.16$, $MSe = 0.10$,

**Table 4-5 Results from the mixed ANOVAS**

**Test of Within-Subjects Effects for Semantic lures**

**(Correct Rejections in spaced repetition type)**

| | Sum of Squares | *df* | Mean Square | F | p |
|---|---|---|---|---|---|
| Lures | 0.078 | 2 | 0.039 | 2.735 | 0.071 |
| Lures*Group | 0.407 | 4 | 0.102 | 7.157 | 0.000* |
| Error (Lures) | 1.194 | 84 | 0.014 | | |
| Times | 0.409 | 2 | 0.204 | 10.101 | 0.000* |
| Times*Group | 0.034 | 4 | 0.009 | 0.422 | 0.792 |
| Error (Times) | 1.699 | 84 | 0.020 | | |
| Lures * Times | 0.398 | 4 | 0.100 | 4.893 | 0.001* |
| Lures*Times*Group | 0.049 | 8 | 0.006 | 0.299 | 0.966 |
| Error (Lures*Times) | 3.412 | 168 | 0.020 | | |

Note: *: $p \leq 0.05$.

**Table 4-6 Results from the mixed ANOVAS**

**Test of Within-Subjects Effects for Semantic lures**

**(Correct rejections in massed repetition type)**

| | Sum of Squares | *df* | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Lures | 0.153 | 2 | 0.077 | 3.491 | 0.035* |
| Lures*Group | 0.065 | 4 | 0.016 | 0.739 | 0.568 |
| Error (Lures) | 1.844 | 84 | 0.022 | | |
| Times | 0.286 | 2 | 0.143 | 0.894 | 0.000* |
| Times*Group | 0.841 | 4 | 0.210 | 13.157 | 0.000* |
| Error (Times) | 1.342 | 84 | 0.016 | | |
| Lures * Times | 0.711 | 4 | 0.193 | 9.486 | 0.000* |
| Lures*Times*Group | 2.047 | 8 | 0.256 | 12.588 | 0.000* |
| Error (Lures*Times) | 3.416 | 168 | 0.020 | | |

Note: *: p≤0.05.

Moreover, grammatical gender effects mentioned above occurred only in the group of "masc.-neu." lures. That is, the proportion of correct rejections to feminine lures (gender-unrelated lures) was significantly higher than that to both masculine and neuter lures (gender-related lures), $F(1, 264) = 5.92$, $MSe = 0.27$. In contrast, the other gender-unrelated lures (neuter lures in the fem.-masc. group and masculine lures in the fem.-neu. lures) led to no more correct rejections than gender-related lures (feminine and masculine lures in the fem.-masc. group and feminine and neuter lures in the fem.-masc. group), $F < 1$ (Figure 4-1, details are in Appendix E - 1).

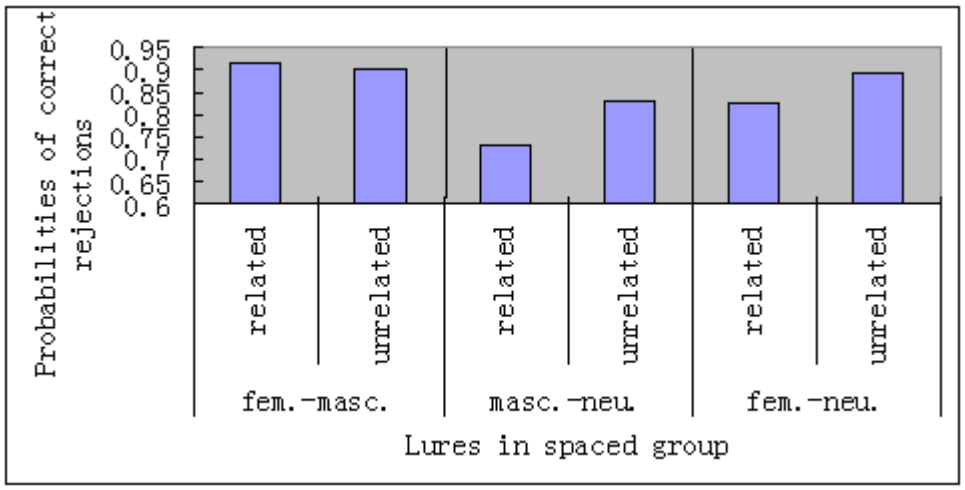



**Figure 4-1 Comparisons between correct rejections of gender-related lures and gender-unrelated lures in spaced group**[7], [8]

Furthermore, comparisons were done between proportions of gender-related lures in the same group (Figure 4-2, details are in Appendix E - 2). Results were influenced by "gender of lures", "number of repetitions", "types of repetition", and "genders combination groups", respectively (Table 4-4). Therefore, gender-related lures in the same combination group were compared separately according to the different factors mentioned above. This comparison revealed fewer correct rejections to feminine than masculine lures in group "fem.-masc.", t (44) = 2.65, $p$ = 0.02, and more correct rejections to feminine than neuter lures in group "fem.-neu.", t (44) = 2.26, $p$ = 0.04. Such differences occurred only when the targets were repeated four times and were

[7] Gender-related: data gender-related words in the "fem.-masc." group was mean of feminine and masculine, in the "masc.-neu." group was mean of masculine and neuter, in "fem.-neu." group was mean of feminine and neuter.

[8] Gender-unrelated words are neuter in "fem.-masc." group or feminine in "masc.-neu." group, or masculine in "fem.-neu." group.

spaced in the learning stage; however, when the targets were presented massedly, a difference between related lures was not found.

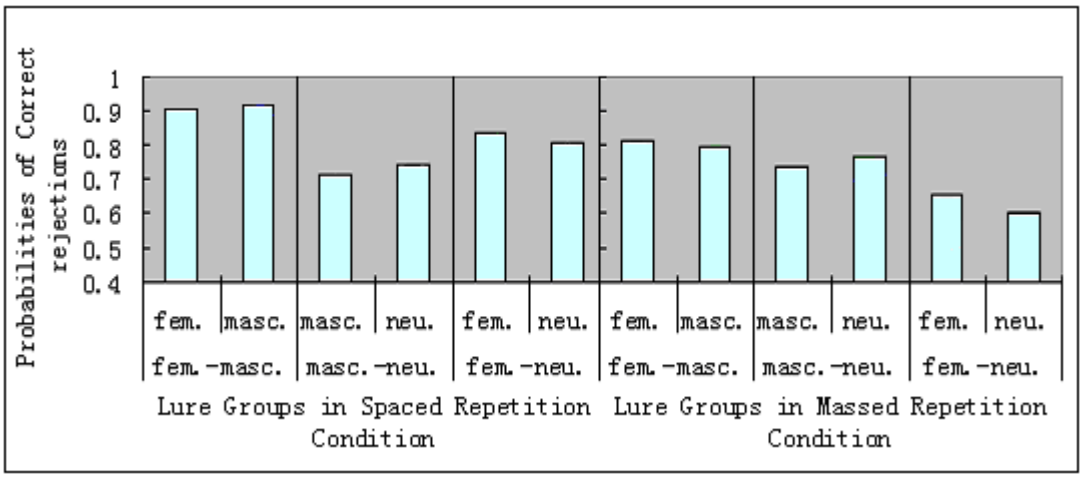



**Figure 4-2 Comparisons between correct rejections of gender-related lures with different repetition types**

Secondly, across factors of number of repetitions, and gender combination groups, feminine lures caused much more correct rejections than masculine and neuter lures, F (2, 168) =4.95, *MSe* = 0.09 (Figure 4-3), especially when the targets were repeated massedly, F (2, 84) = 3.49, *MSe* = 0.08. By contrast, such difference were not significant at the alpha level of .05, F (2, 84) = 2.74, *MSe* = 0.04, when the targets were repeated spacedly.

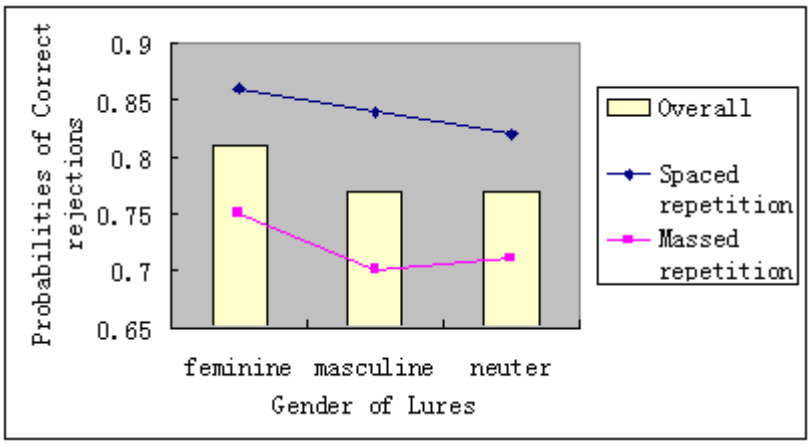



**Figure 4-3 Comparisons of correct rejections among different lures (across different number of repetitions and different target groups)**

Thirdly, a main effect of gender combination groups was found: different combination groups caused different proportions of correct rejection, F (2, 84) = 5.63, *MSe* = 1.10. In addition, interaction between types of repetition and combination groups was found, F (2, 84) = 4.65, *MSe* = 0.91. It should be also noted that proportions of correct rejections to group "fem.-neu." with massed repetition were extremely low. Half of the data were near or lower than 0.5, although materials used in this type were the same used in spaced repetition type.

Finally, a main effect of the number of repetitions was also found, F (2, 168) = 11.81, *MSe* = 0.21; proportions of correct rejections decreased mainly along with the increase of the number of repetitions: Multiple comparisons showed that repeating one time caused more correct rejections than repeating more times ($p$ = 0.00), while proportion of correct rejections did not change significantly when the number of repetitions was increased from 4 times to 8 times ($p$ = 0.82). Similar time effects were found both with spaced repetition type and with massed repetition type.

#### 4. 2. 2. 1. 2. False Alarms to the Lures

Similar analyses on the false alarms to semantic related lures, including gender-related and gender-unrelated lures, were done. It showed that the proportions of false "Know" and "Guess" were not influenced by either different number of repetitions, $F$ (2, 168) ≤ 1.30, or different gender combination groups, $F$ (2, 84) ≤ 2.58. There was also no interaction between gender of lures and gender combination groups either with false "Know", $F$ (4, 168) = 1.73, $MSe$ = 0.01, or false "Guess", $F$ (4, 168) = 1.55, $MSe$ = 0.02. However, massed repetition caused more false "Know" than spaced repetition, F (1, 84) =11.82, $MSe$=0.24, but the same effect was not found with false "Guess", $F$ (1, 84) < 1 (Figure 4-4).

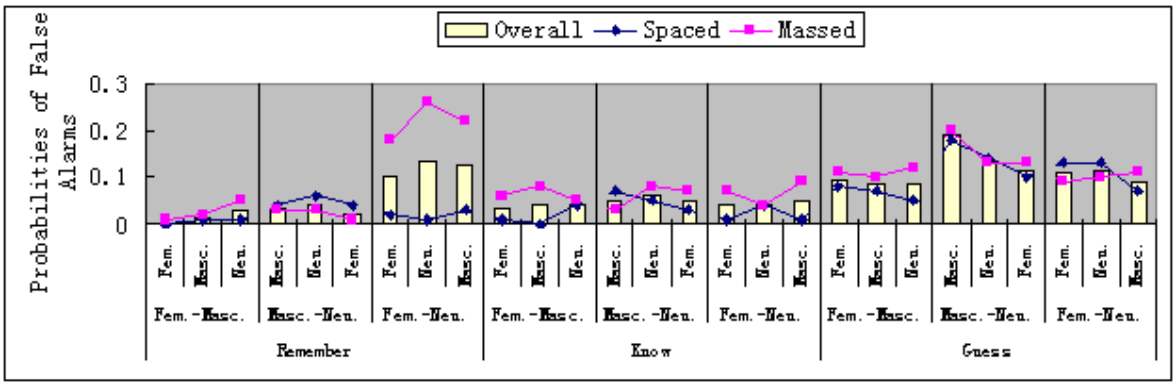



**Figure 4-4 Comparisons of false alarms among different genders, repetition types and response types**

By contrast, proportions of false "Remembers" were affected by an interaction between "types of repetition" and "gender combination groups". False alarms to lures in the group "fem.-neu," were significant more than those in either group "fem.-

masc." or group "masc.-neu.", when the learned targets were massedly repeated, F (2, 42) = 137.84, *MSe* = 1.70. In addition, false "Remembers" were influenced by the number of repetitions, F (2, 168) = 8.13, MSe = 0.03. The more the number of repetitions, the higher the proportions of false "Remembers" (details are on Appendix B - 4).

#### 4. 2. 2. 2. Analysis of Targets (Hits)

Since the materials in each gender combination group (group "fem.-masc.", group "masc.-neu.", and group "fem.-neu.") were not the same, all data were analyzed separately according to the gender combination groups, with comparisons between proportions of hits to feminine and masculine targets in group "fem.-masc.", proportions of hits to masculine and neuter targets in group "masc.-neu.", and proportions of hits to feminine and neuter targets in group "fem.-neu," conducted. Statistical analyses consisted of separate 2 (genders of targets: feminine and masculine in group "fem.-masc.", or masculine and neuter in group "masc.-neu.", or feminine and neuter in group "fem.-neu.") × 3 (number of repetitions: words were repeated once, four, or eight times) × 2 (types of repetition: spaced or massed) analyses of variance (ANOVAs) (Appendix B - 5).

One of the main issues in this experiment was to detect whether there are differences among remembering and identifying words which belong to different genders. This comparison showed that, across the repetition type, no difference was found with hits "Remembers": hits "Remembers" of feminine targets were not

significantly greater than that masculine targets in group "fem.-masc.", $F$ (1, 88) = 3.59, $MSe$ = 0.08; and were not significantly less than neuter targets in group "fem.-neu." $F$ (1, 88) = 1.70, $MSe$ = 0.05; in group "masc.-neu.", no significant difference was found between proportions of hits "Remember" of masculine and neuter targets, either, $F$ (1, 88) = 2.93, $MSe$ = 0.08. The same results were found also with hits "Know", $F < 1$. By contrast, results showed significant difference only between proportions of hits "Guess" of feminine and masculine targets in group "fem.-masc." $F$ (1, 88) = 6.11, $MSe$ = 0.07.

In addition, across different genders (feminine, masculine or neuter) and different gender combination groups (group "fem.-masc.", group "masc.-neu.", or group "fem.-neu"), result revealed that proportions of hits "Remember" were influenced by number of repetitions, F (2, 356) = 136.17, $MSe$ = 4.28. The more the words were repeated, the more successfully participants remembered the target words. Such a number-of-repetition effect influenced spaced repetition more than massed repetition, F (2, 356) = 29.20, $MSe$ = 0.92. Moreover, spaced repetition caused more hits Remember than massed repetition, F (1, 178) = 7.14, $MSe$ = 0.93.

In contrast to hits of "Remember", main effect of repetition type was not found with hits of "Know", $F < 1$. Hits of "Know" were influenced mainly by effect of number of repetitions, F (2, 356) = 10.73, $MSe$ = 0.22; while hits of "Guess" were affected by types of repetition, as well as effects from number of repetitions, F (2, 356) = 4.68, $MSe$ = 0.06. Targets words which were repeated in a spaced fashion were marked as "guesses" more often than target words that were repeated in a massed

fashion, F (1, 178) = 13.25, *MSe* = 0.38.

### 4. 2. 3. Discussion

In Experiment 2, grammatical gender effects were indicated by the interaction between "gender of lures" and "gender combination groups". Results revealed the expected grammatical gender effects in group "masc.-neu.". Feminine lures, which were not included in the learning stage, had more correct rejections than masculine and neuter lures. Similar effects were not found in group "fem.-masc." or in group "fem.-neu.". Neither masculine nor neuter words, which were arranged as the gender-unrelated lures in group "fem.-masc." and group "fem.-neu." respectively, caused higher proportions of correct rejections than gender-related lures. Such results indicate that feminine lures can be more easily correct rejected when it was not learned, although it was influenced by the same semantic category effect as masculine and neuter lures were, which was revealed to be one of the main factors that affect recognition tasks. By contrast, masculine and neuter, either of which was gender unrelated in group "fem.-neu," and group "fem.-masc." respectively, were influenced less by the grammatical gender effects according to the results. However, masculine lures, as gender-unrelated lures, were more easily rejected correctly than feminine lures (gender-related lures) in Experiment 1, it is claimed that the gender information of masculine and neuter words was less affected by grammatical gender effects only when one of them was gender related while the other one was gender unrelated. This conclusion was also indicated by comparing correct responses of masculine words

and neuter words in the same group when they were both related lures/targets. No differences were found between identifying masculine and neuter words.

Moreover, differences were found between identifying either feminine words and masculine words or feminine words and neuter words. However, such differences were found only when targets were repeated four times. It seems that these differences were not influenced by the increase of number of repetitions when the materials were displayed without explicit gender information (bare nouns).

Another result concerns semantic category effects. By analyzing targets, the proportion of hits "Remember" increased with the number of times an item was presented. Such number-of-repetition effects influenced spaced repetition more than massed repetition. In addition, spaced repetition led to more hits than massed repetition. In contrast to hits of "Remember", hits of "Know" varied not only according to the alteration of number of repetitions, but also by different types of repetition, while hits of "Guess" were found less frequently in spaced repetition type than that in massed repetition type.

Finally, as a whole, spaced repetition enhanced the grammatical gender effects. Difference between proportions of gender-related lures and gender-unrelated lures was found only in the spaced repetition type, but not in massed repetition type. However, massed repetition enhanced the difference in identifying different genders. Feminine lures caused more correct rejections than masculine and neuter lures only when targets were arranged by massed repetition. In short, semantic category effects were similar to what was reported in the recent studies.

# 5. Experiment 3

In the following experiment, definite determiners noun phrases were used as the materials, so that grammatical gender information could be explicitly presented. In German, there are three definite determiners applied according to the gender of the noun (i.e. "die" for feminine nouns [die Katze – the cat], "der" for masculine nouns [der Hund – the dog], and "das" is assigned to neuter nouns [das Pferd – the horse]). Since the determiners, including gender-marked adjectives, might influence the noun phrase production (Levelt, et al. 1999; Joël Macoir and Renée Béland, 2004; Roelofs, 1992), it is claimed that determiners might affect the recognition task, as well. Namely, the recognition task might be either facilitated or inhibited by such an explicit gender cue.

As the results in Experiment 2 indicated, grammatical gender effects were found only when masculine and neuter were arranged as related genders while feminine was arranged as an unrelated gender (group "masc.-neu."). In addition, when correct rejections of related lures were compared, significant differences occurred between feminine lures and masculine lures in group "fem.-masc.", and between feminine lures and masculine lures in group "fem.-neu." when the targets were repeated four times. Such a result indicated a facilitation of gender cue when feminine words were used only in the recognition task. Moreover, when feminine was displayed with either masculine or neuter in the learning stage, gender information was identified and one of the genders was better remembered than the other one. However, such an effect was not strong enough with bare nouns. These suppositions will be clarified in

Experiment 3 with determiner noun phrases.

In Experiment 3, determiner noun phrases were used instead of bare nouns, so that the gender information of words could be explicitly displayed. Determiners are likely to facilitate the recognition memory, so that gender-unrelated lures would have higher proportions of correct rejections than gender-related lures in the recognition task, in the "fem.-masc." and "fem.-neu." groups as well as the "masc.-neu." group. In addition, the difference between the proportions of correct rejections to masculine and neuter lures in group "masc.-neu." should become significant. Otherwise, the determiner-noun-form can bee seen to inhibit the recognition tasks, if the grammatical gender effects found in Experiment 2 vanish.

## 5. 1. Method

**Participants:** 45[9] native German speakers (15 male, 30 female, mean age 23.29, *SD*=3.51) from Heidelberg University were involved in this experiment.

**Stimuli:** The nouns used in Experiment 2 were used again in this experiment. The arrangement of materials was nearly the same as in Experiment 2. That is, two items, feminine and masculine, feminine and neuter, or masculine and neuter, were selected from each category as the targets, which were presented in the learning phase and used as the targets in the subsequent recognition test. Three items, including 1 feminine, 1 masculine and 1 neuter noun, were selected from the remaining four words in each semantic category and were used as the semantically related lures.

[9] The sample size of 45 yields a statistical power of .9990, given an effect size of .25 and an alpha-level of .05 (analysis based on G*Power; see Faul et al., 2007).

Therefore, there were two kinds of semantically related lures: gender-related lures and gender-unrelated lures. Also, 18 filler items were set up as semantically unrelated lures that did not share the same semantic category with the targets. All words were displayed with appropriate determiners. Lures, including gender-related lures, gender-unrelated lures, and filler items, were only displayed in the recognition stage. Frequencies of words were controlled. Words used as targets, gender-related lures, and gender-unrelated lures in each semantic category were tested to be semantically close to each other in the pretest in Experiment 2.

In the previous three-gender condition study, ultra data were found when target words were presented massed and the tendency of the correct rejection caused by the grammatical gender effects was not significant; therefore, in this experiment, only spaced repetition was used as the repetition type. Appropriate determiners were presented with all nouns in both the learning stage and the recognition task, for example *der Elefant*, *das Lamm*, etc. That is, not bare nouns but noun phrases were used as materials.

**Procedure**: Procedures were similar to those used in Experiment 1. Study items were presented on Apple Macintosh computers. Each word remained on the screen for one second and was replaced after an interval of one second by the next item. Participants were instructed to read the words silently as they appeared on the screen and to bear in mind that they would later be given some form of memory test, the precise nature of which was not specified. In the recognition task, participants were asked to identify the words they had learned with answers "*Remember*", "*Know*", "*Guess*", or "*No*"

(details are in Experiment 1). Between the learning stage and the recognition task, a distractor task named *Rushhour* was provided for 15 minutes.

## 5. 2. Results

As in Experiment 2, this analysis focused on the difference between the correct rejections of gender-related lures and gender-unrelated lures. The main results, means of hits, correct rejections of gender-related lures and gender-unrelated lures can be seen in Table 5-1 and Table 5-2 (more details are in Appendix C - 1). An alpha level of .05 was used in all statistical analyses.

**Table 5-1 Response Proportions (*p*) as a Function of Exact Repetition for Targets**

| | | Target group "fem.-masc." | | | Target group "masc.-neu." | | | Target group "fem.-neu." | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| | R* | 0.47 | 0.86 | 0.86 | / | / | / | 0.27 | 0.70 | 0.90 |
| F | K | 0.07 | 0.11 | 0.04 | / | / | / | 0.30 | 0.20 | 0.03 |
| | G | 0.18 | 0.00 | 0.10 | / | / | / | 0.17 | 0.03 | 0.07 |
| | R | 0.36 | 0.67 | 0.85 | 0.38 | 0.69 | 0.82 | / | / | / |
| M | K | 0.07 | 0.09 | 0.13 | 0.07 | 0.09 | 0.09 | / | / | / |
| | G | 0.16 | 0.17 | 0.02 | 0.07 | 0.09 | 0.05 | / | / | / |
| | R | / | / | / | 0.40 | 0.69 | 0.67 | 0.44 | 0.70 | 0.70 |
| N | K | / | / | / | 0.06 | 0.11 | 0.09 | 0.13 | 0.10 | 0.03 |
| | G | / | / | / | 0.09 | 0.09 | 0.13 | 0.13 | 0.10 | 0.20 |

Note: F: feminine targets, M: masculine targets, N neuter targets.

R, K, G denotes answers Remember, Know, and Guess respectively.

**Table 5-2 Response Proportions (*p*) as a Function of Exact Repetition for Lures (Correct Rejections only)**

| | Lure group "fem.-masc." | | | Lure group "masc.-neu." | | | Lure group "fem.-neu." | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| F | 0.93 | 0.88 | 0.83 | 0.78 *u* | 0.83 *u* | 0.73 *u* | 0.80 | 0.80 | 0.77 |
| M | 0.81 | 0.78 | 0.67 | 0.82 | 0.91 | 0.83 | 0.93 *u* | 0.93 *u* | 0.63 *u* |
| N | 0.96 *u* | 0.65 *u* | 0.78 *u* | 0.89 | 0.75 | 0.70 | 0.97 | 0.93 | 0.70 |

Note: F: feminine lures, M: masculine lures, N neuter lures.

Data with *u* were proportions of correct rejections to grammatical gender-unrelated lures.

### 5. 2. 1. Correct Rejections of Gender-Related and Gender-Rnrelated Lures

Correct rejections to the lures were analyzed. Since "gender combination group" (groups "fem.-masc.", "masc.-neu.", and "fem.-neu.") was a between-subject factor, statistical analyses included 3 (genders of lures: feminine, masculine, neuter) × 3 (gender combination groups: "fem. - masc.", "masc. - neu.", and "fem. - neu.") × 3 (number of repetitions: one, four, eight times) mixed analyses of variance (ANOVAs).

An alpha level of .05 was used in all statistical analyses.

**Table 5-3 Results from the mixed ANOVAS**

**Test of Within-Subjects Effects for Semantic lures (Correct Rejections only)**

| | Sum of Squares | *df* | Mean Square | F | p |
|---|---|---|---|---|---|
| Lures | 0.002 | 2 | 0.001 | 0.050 | 0.951 |
| Lures*Group | 0.709 | 4 | 0.177 | 10.797 | 0.000* |
| Error (Lures) | 1.380 | 84 | 0.016 | | |
| Times | 1.332 | 2 | 0.666 | 27.617 | 0.000* |
| Times*Group | 0.532 | 4 | 0.133 | 5.515 | 0.001* |
| Error (Times) | 2.025 | 84 | 0.024 | | |
| Lures * Times | 0.586 | 4 | 0.147 | 8.627 | 0.000* |
| Lures*Times*Group | 0.490 | 8 | 0.061 | 3.608 | 0.001* |
| Error (Lures*Times) | 2.854 | 168 | 0.017 | | |

Note: *: $p \leq 0.05$.

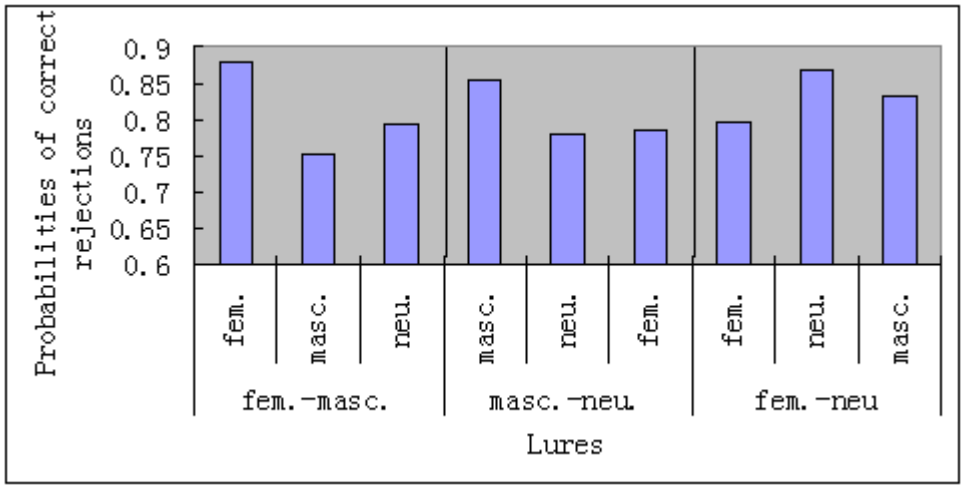



**Figure 5-1 Comparisons of correct rejections among different lures in different gender combination groups**

Significant interaction was found between "gender of lures" and "gender combination group", F (4, 84) = 10.80, *MS*e = 0.18. Such an interaction was not

consistent with the main expectation in this experiment. The main expectation was that gender-unrelated lures would have higher proportions of correct rejections than gender-related lures in the recognition task in not only group "masc.-neu.", but also group "fem.-masc." and "fem.-neu.". However, results showed no significant difference between gender-related lures and gender-unrelated lures in either group "fem.-masc.", t (14) = 0.91, $p$ = 0.38, or in group "fem.-neu.", t (14) = 0.31, $p$=0.76; in addition, the difference between gender-related lures and gender-unrelated lures in group "masc.-neu.", t (14) =1.53, $p$=0.15, was not significant either (Figure 5-1, details are in Appendix E - 3). By contrast, when related lures were compared, feminine lures caused more correct rejections than masculine lures in group "fem.-masc.", t (14) = 4.49, $p < 0.01$, but less than neuter lures when they were in group "fem.-neu.", t (14) =2.35, $p$=0.03. In addition, the correct rejection of masculine lures was much higher than that of neuter lures in group "masc.-neu.", t (14) = 3.74, $p < 0.01$.

In addition, the proportions of correct rejections to the gender-unrelated lures and the gender-related lures were compared. The gender-related lures had fewer correct rejections in each gender combination group. Results showed that the correct rejection of neuter lures was not significant different from masculine lures in group "fem.-neu", t (44) = 1.44, $p$ = 0.16; differences between neuter lures and feminine lures in group "masc.-neu.", t (44) = 1.36, $p$ = 0.18 and between feminine lures and masculine lures in group "fem.-neu.", t (44) = 0.40, $p$ = 0.69, were not found either. Across different gender combination groups, differences among proportions of correct rejections to feminine, masculine and neuter lures were not significant ($F < 1$).

Moreover, main effect of number of repetitions was again found, $F$ (2, 84) = 27.62, $MSe$ = 0.67. A further contrast showed that proportions of correct rejections decreased along with the increasing of number of repetitions ($M_1$ = 0.88, $M_4$ = 0.83, $M_8$ = 0.74, $p$s ≤ 0.01). Such an effect was similar to the result in Experiment 2.

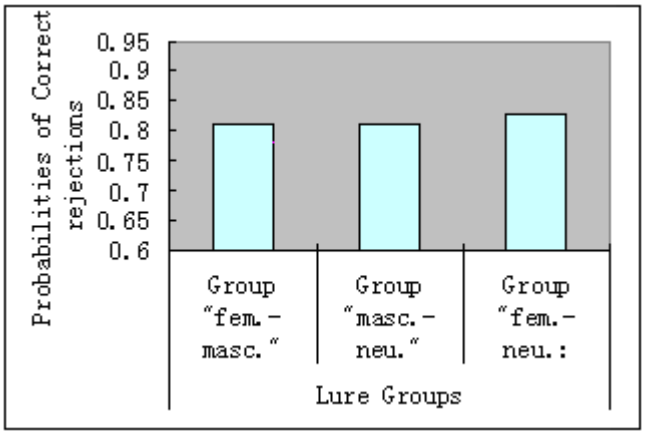



**Figure 5-2 Comparisons of correct rejections among different gender combination groups (Across different genders)**

It should be also noted that differences among different gender combination groups (group "fem.-masc.", group "masc.-neu.", and group "fem.-neu.") were not significant $F$ < 1 (figure 5-2, details are in Appendix E - 4). This result was different from that in Experiment 2. In Experiment 2, correct rejection occurred in the group "fem. - masc." lures was significantly more than that in the group "masc. - neu." lures.

### 5. 2. 2. False Alarms to Gender-Related and Gender-Unrelated Lures

All "Remember", "Know", and "Guess" responses to the lures were false alarms. Similar analyses on the false alarms to lures, including gender-related and gender-unrelated lures, were applied.

Interaction between "gender of lures" and "gender combination group" existed not only with the answer "Remember", F (4, 84) = 8.39, *MSe* = 0.02, but also with answer "Know", F (4, 84) = 4.45, *MSe* = 0.02, and with answer "Guess", F (4, 84) = 6.09, *MSe* = 0.12. However, these interactions were apparently not caused by differences between grammatical gender-related lures and unrelated lures (Figure 5-3). As seen with correct rejections, one of the gender-related lures had more false alarms than the other when proportions of false alarms to gender-related lures were compared (i.e. feminine v.s. masculine in group "fem.-masc.", masculine v.s. neuter in group "masc.-neu.", and feminine v.s. neuter in group "fem.-neu.").

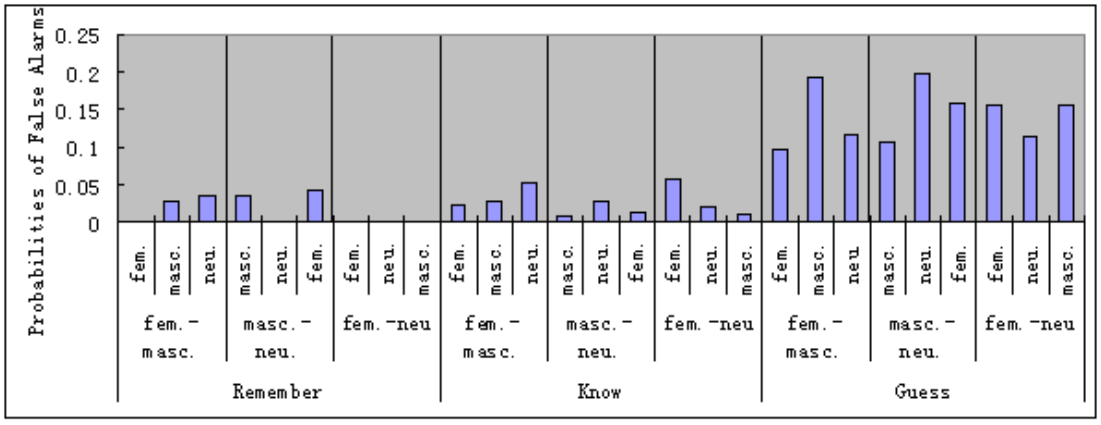



**Figure 5-3 Comparisons of false alarms among different lures within each gender combination group with different response type**

Moreover, false "Remember" responses, $F$ (2, 84) = 19.84, MSe = 0.05, and false "Guess" responses to lures, $F$ (2, 84) = 17.00, MSe = 0.29, increased along with increasing of the repetition times of the learning targets; however, the rate of false "Know" responses was not influenced by alteration of number of repetitions, $F$(2, 84) = 1.78, MSe = 0.01, $p$ = 0.17.

### 5. 2. 3. Response Proportions of Targets (Hits)

Since materials used in different gender combination groups (group "fem.-masc.", group "fem-neu.", and group "masc.-neu.") were not the same and cannot be combined together, all target responses pertaining to different groups were analyzed separately. Within each gender combination group, statistical analyses included 2 (genders of targets: feminine and masculine in group "fem. - masc.", or masculine and neuter in group "masc. - neu.", or feminine and neuter in group "fem. - neu.") × 3 (number of repetitions: one, four, or eight times) mixed analyses of variance (ANOVAs) (Appendix C - 2).

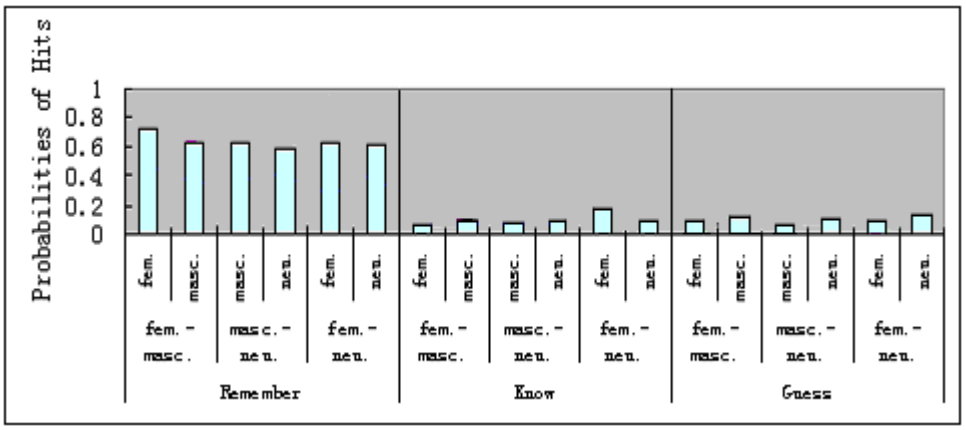



**Figure 5-4 Comparisons of hits between targets within each gender combination group with different response type**

In group "masc.-neu.", the proportion of not only hits of "Remember", but also of "Know" and "Guess" showed no significant difference between masculine targets and neuter targets, *F*s < 1.18. In contrast to group "masc.-neu.", feminine noun phrases caused more hits of "Remember" than masculine noun phrases did in group "fem.-masc.", $F(1, 14) = 22.88$, $MSe = 0.24$, but they caused fewer hits of "Know" than neuter noun phrases in group "fem.-neu." did, $F(1, 14) = 7.55$, $MSe = 0.18$ (Figure 5-4, details are in Appendix E - 5).

Furthermore, in this experiment, the proportion of hits to the targets was again found to increase with the increase of number of repetitions with answer "Remember", $F(2, 178) = 146.54$, $MSe = 4.32$, but decreased with the increase of repetition with respect to answer "Know", $F(2, 178) = 3.09$, $MSe = 0.07$; however, the proportion of answer "Guess" was also influenced by the amount of repetition, $F(2, 178) = 3.73$, $MSe = 0.06$. The more times the targets were repeated, the fewer hits of "Guess" occurred.

## 5. 3. Discussion

In this experiment, it was assumed that participants would correctly reject gender-unrelated lures more frequently than gender-related lures. However, such a grammatical gender effect was not found as expected. By contrast, as gender-related lures, feminine lures caused more correct rejections than masculine lures in the group "fem.-masc.", but fewer correct rejections than neuter lures in the group "fem.-neu.".

At the same time, the proportion of correct rejection of masculine lures was

higher than neuter lures in group “masc.-neu.”. Such difference was called “gender conflict”. Data showed that, with the “gender conflict”, one of the related genders has been facilitated to be remembered in the learning stage and used as the gender cue in the following recognition task. The better the gender cue was used, the more correct rejections occurred. The opportunity of facilitation was equal to each gender. None of the genders can consistently cause more correct rejections than the others. It should be noted that, this difference between masculine lures and neuter lures found in Experiment 3 was not found in Experiment 2, in which only bare nouns were used. It indicated that the definite determiners enhanced at least the gender information of masculine lures, so that the proportion of correct rejection to masculine lures versus neuter lures was significantly higher in this experiment. Moreover, although the gender conflict occurred between masculine lures and neuter lures in group “masc.-neu.”, a significant difference of proportions was not found between correct identification of masculine targets and of neuter targets.

In addition, results from correct identification indicated that the semantic category effect was found to occur with determiner noun phrases. The proportion of hits was mainly influenced by the variation of the number of repetitions.

In short, different results were found in Experiment 2 and 3. That is: (1) significant grammatical gender effects were found in the recognition task with bare nouns. Correct rejections of the gender-unrelated lures were more frequent than correct rejections of the gender-related lures in group “masc.-neu.”; however, with the definite determiners included in the noun phrase, no significant difference was found

with in the three different gender combination groups; (2) Gender conflict was found between masculine lures and neuter lures in group “masc.-neu.”, only when definite determiners noun phrases, but not bare nouns, were used as experiment materials.

The different results between Experiment 2 and 3 described above might be caused by different determiner conditions. Without determiners, grammatical gender information carried by bare nouns affects the recognition tasks implicitly. In such a situation, feminine lures were more easily identified as gender-unrelated lures (in group “masc.-neu.”) than either masculine lures or neuter lures (in group “fem.-masc.” and “fem.-neu.”, respectively) and gender conflict was not found in group “masc.-neu.”. Feminine seemed to be the easiest gender for participants to identify. However, definite determiners noun phrases decreased the grammatical gender effects and enhanced the conflict between identifying masculine lures and neuter lures, both of which were gender-related lures, in group “masc.-neu.”. Therefore, it is assumed that the gender information of masculine and neuter nouns was inhibited so that feminine can be more often correctly identified only when explicit gender cues (e.g. definite determiners) are missing. By contrast, once there were the explicit gender cues (Experiment 3), and both gender information and semantic information were displayed, recognition tasks became so complicated that the advantage of the feminine gender in being more often correctly/falsely remembered vanished. However, these hypotheses were not yet confirmed with the experiments above. In the following experiment, the remaining issues were addressed.

# 6. Experiment 4

In this experiment, the words used as materials were displayed with pseudo-determiners. In the previous experiments, I focused on participants' recognition of bare nouns and noun phrases with definite determiners, and found that grammatical gender showed different influences on recognition tasks with different materials (bare nouns or definite determiner noun phrases) when repeated word lists were used as materials.

However, in the daily speaking, the endings of either determiners or adjectives used with masculine and neuter words are frequently the same. That is, indefinite determiners and negative determiners used either with masculine words or with neuter words (nominative) have no endings (indefinite determiner *ein* - a, negative determiner *kein* - no), and ending -em is used by dative adjectives for both masculine and neuter words, when neither definite determiners nor indefinite determiners are used; by contrast, both indefinite determiner (*eine* - a) and negative determiners (*keine* - no) used with feminine words (nominative) have the ending -e; and dative adjectives for feminine words have the ending -er. Previously, studies on lexical access showed that transparent gender suffix of articles facilitated the gender judgment tasks (Taft et al., 1998). Therefore, this experiment investigated the influence of gender suffixes used in daily speaking on recognition tasks. Since no indefinite determiners are used with uncountable nouns, and the gender suffixes of either determiners or adjectives were the main focus, pseudo-determiners "elm" and "elme", which showed the transparent ending difference, were used instead of the real determiners. Namely,

"elm" was used instead of articles for masculine and neuter nouns, and "elme" was used instead of articles for feminine nouns.

There were three main issues under investigation: (1) grammatical gender effects in my research were represented in the difference between proportions of correct rejection of gender-related lures and gender-unrelated lures. Why was such an effect found only in Experiment 2? Data showed that definite determiners used in Experiment 3 might make the learning and recognition tasks more complicated so that the grammatical gender effects vanished. If grammatical gender effects were not found in Experiment 4 either, determiner-noun-phrases did inhibit the recognition task; (2) each participant received two kinds of gender-related lures in the recognition task (e.g. participants received feminine lures and masculine lures after learning feminine targets and masculine targets) in the previous studies. After these experiments, it was still unclear why significant differences occurred between proportions of correct rejection of different lures. It is assumed that one gender can be kept in mind more strongly than the other one when people try to remember two kinds of words which belong to two different genders. And the determiners might be the explicit signs which make the difference become significant. In Experiment 4, since masculine words and neuter words shared the same pseudo determiner "elm", the above hypothesis can be confirmed if the conflicts between gender-related lures in group "masc.-neu.", which were found in Experiment 3, disappeared; (3) The previous experiments did not resolve whether masculine and neuter noun phrases influenced the recognition task similarly while the feminine words did not. It was assumed that

characteristics of masculine and neuter might become similar when masculine and neuter words share the same determiner (*ein*) in the daily speaking; by contrast, characteristics of feminine words might become more noticeable. In Experiment 4, masculine and neuter words were displayed with the same pseudo-determiner (*elm*), while feminine words were presented with a different one (*elme*). Masculine and neuter noun phrases will be shown to influence recognition task similarly if grammatical gender effects, which are similar to that in Experiment 2, again occur in Experiment 4.

## 6. 1. Method

**Participants:** 45[10] native German speakers from Heidelberg University (17 male, 28 female, mean age 22.84, *SD*=3.74) participated in this research.

**Stimuli:** The nouns used in Experiment 3 were used again in this experiment. Arrangement of materials was nearly the same as in Experiment 3. That is, two items, either feminine and masculine, feminine and neuter, or masculine and neuter, were selected from each category as the targets, which were presented in the learning phase and used as the targets in the subsequent recognition test. Three items, including 1 feminine, 1 masculine and 1 neuter noun, selected from the remaining four words in each semantic category, were used as the semantically related lures. Therefore, there were two kinds of semantically related lures: gender-related lures and gender-unrelated lures. Also, 18 filler items were used as semantically unrelated lures that did

[10] The sample size of 45 yields a statistical power of .9990, given an effect size of .25 and an alpha-level of .05 (analysis based on G*Power; see Faul et al., 2007).

not share the same semantic category with the targets. All lures, including gender-related lures, gender-unrelated lures, and filler items, were only displayed in the recognition stage. The frequency of words was controlled. Words used as targets, gender-related lures, and gender-unrelated lures in each semantic category were tested to be semantically close to each other in the pretest in Experiment 2.

As in Experiment 3, only spaced repetition was used as the repetition type. Pseudo-determiners were presented with all nouns in both learning stage and recognition task in this experiment, for example *elm Elefant* (masc.), *elm Lamm* (neu.), *elme Katze* (fem.) etc. That is, masculine nouns and neuter nouns shared the same pseudo-determiners "*elm*", while feminine nouns occurred with "*elem*".

**Procedure**: Procedures were similar to that used in Experiment 1. Study items were presented on Apple Macintosh computers. Each word remained on the screen for one second and was replaced after an interval of one second by the next item. Participants were instructed to read the words silently as they appeared on the screen and to bear in mind that they would later be given some form of memory test, the precise nature of which was not specified. In the recognition task, participants were asked to identify the words they had learned with answers "*Remember*", "*Know*", "*Guess*", or "*No*" (details are in Experiment 1). Between the learning stage and the recognition task, a distractor task named *Rushhour* was provided for 15 minutes.

## 6. 2. Results

The analysis again focused primarily on the difference between correct rejections of gender-related lures and gender-unrelated lures. The main results, means of hits, and correct rejections of gender-related lures and gender-unrelated lures can be seen in Table 6-1 and Table 6-2 (more details are in Appendix D - 1). All "Remember" responses (chosen if participants can consciously recollect seeing the item at study), "Know" responses (chosen if the item seems familiar but participants cannot recollect its earlier presentation), and "Guess" responses (chosen for items that participants neither recollect nor recognize on the basis of familiarity, but which they cannot definitely reject) to the lures are false alarms, while the rest, "*No*" responses (participants believe that an item did not appear in the earlier phase), are correct rejections.

**Table 6-1 Response Proportions (*p*) as a Function of Exact Repetition for Targets**

| | | "fem.-masc." group targets | | | "masc.-neu." group targets | | | "fem.-neu." group targets | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| | R* | 0.38 | 0.83 | 0.71 | / | / | / | 0.44 | 0.73 | 0.84 |
| F | K | 0.12 | 0.09 | 0.18 | / | / | / | 0.09 | 0.13 | 0.08 |
| | G | 0.10 | 0.05 | 0.03 | / | / | / | 0.17 | 0.08 | 0.05 |
| | R | 0.34 | 0.65 | 0.74 | 0.36 | 0.63 | 0.86 | / | / | / |
| M | K | 0.22 | 0.18 | 0.09 | 0.19 | 0.21 | 0.13 | / | / | / |
| | G | 0.16 | 0.09 | 0.12 | 0.05 | 0.07 | 0.01 | / | / | / |
| | R | / | / | / | 0.33 | 0.77 | 0.81 | 0.41 | 0.73 | 0.77 |
| N | K | / | / | / | 0.18 | 0.12 | 0.16 | 0.08 | 0.06 | 0.06 |
| | G | / | / | / | 0.17 | 0.03 | 0.03 | 0.16 | 0.07 | 0.13 |

Note: F: feminine targets, M: masculine targets, N: neuter targets.

R, K, G denotes answers Remember, Know, and Guess respectively.

**Table 6-2 Response Proportions (*p*) as a Function of Exact Repetition for Lures (Correct Rejections only)**

| | "fem.-masc." group lures | | | "masc.-neu." group lures | | | "fem.-neu." group lures | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| F | 0.87 | 0.91 | 0.77 | 0.87 *u* | 0.83 *u* | 0.89 *u* | 0.74 | 0.72 | 0.64 |
| M | 0.83 | 0.79 | 0.78 | 0.90 | 0.81 | 0.80 | 0.75 | 0.79 | 0.58 |
| N | 0.90 *u* | 0.72 *u* | 0.77 *u* | 0.90 | 0.80 | 0.77 | 0.91 *u* | 0.77 *u* | 0.63 *u* |

Note: F: feminine lures, M: masculine lures, N: neuter lures.

Data with *u* were from grammatical gender-unrelated lures.

### 6. 2. 1. Correct Rejections of Gender-Related and Gender-Unrelated Lures

As in Experiment 3, "gender combination group" (groups "fem. - masc.", "masc. - neu.", and "fem. - neu.") was a between-subject factor. Statistical analyses included 3 (genders of lures: feminine, masculine, neuter) × 3 (gender combination groups: "fem. - masc.", "masc. - neu.", and "fem. - neu.") × 3 (Number of repetitions: one, four, eight times) mixed analyses of variance (ANOVAs).

An alpha level of .05 was used in all statistical analyses.

**Table 6-3 Results from the mixed ANOVAS**

**Test of Within-Subjects Effects for Semantic lures (Correct Rejections only)**

| | Type Ⅲ Sum of Squares | *df* | Mean Squares | F | p |
|---|---|---|---|---|---|
| Lures | 0.041 | 2 | 0.021 | 1.033 | 0.361 |
| Lures*Group | 0.206 | 4 | 0.051 | 2.553 | 0.045* |
| Error (Lures) | 1.692 | 84 | 0.020 | | |
| Times | 0.876 | 2 | 0.438 | 23.459 | 0.000* |
| Times*Group | 0.276 | 4 | 0.069 | 3.698 | 0.008* |
| Error (Times) | 1.569 | 84 | 0.019 | | |
| Lures*Times | 0.272 | 4 | 0.068 | 2.675 | 0.034* |
| Lures*Times*Group | 0.275 | 8 | 0.034 | 1.356 | 0.220 |
| Error (Lures*Times) | 4.262 | 168 | 0.025 | | |

Note: *: $p \leq 0.05$.

The most important finding was a significant interaction between "gender of lures" and "gender combination groups", $F$ (4, 84) = 2.55, $MS$e = 0.05. Figure 6-1 (details are in Appendix E - 6) shows that in group "fem.-masc." and in group "fem.-

neu.", gender-related lures received more correct rejections than gender-unrelated lures, while the proportions of correct rejections in group "masc.-neu." had the opposite tendency. Further t-tests showed no significant differences between gender-related lures and gender-unrelated lures at the alpha level of 0.05 with either group "fem.-masc.", t (14) = 0.81, $p$ = 0.43, or "fem.-neu", t (14) = 1.05, $p$ = 0.31; by contrast, a significant difference between related lures and unrelated lures was found in group "masc.-neu.", t (14) = 2.91, $p$ = 0.01, meaning that more correct rejections to the gender-unrelated lures occurred than to the gender-related lures.

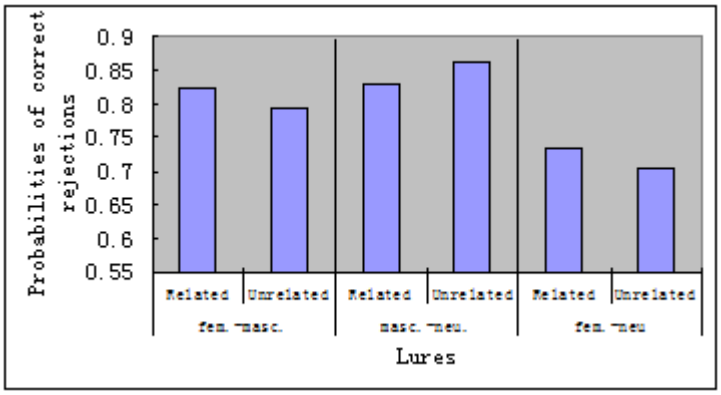



**Figure 6-1 Comparisons between gender-related lures and gender-unrelated lures within each gender combination group (Data across Times)**

By analyzing the correct rejections of gender-related lures (Figure 6-2, details are in Appendix E - 7), across different number of repetitions, dependent t-tests showed significant differences of proportions between feminine lures and masculine lures in group "fem.-masc.", t (14) = 2.24, $p$ = 0.04, and between feminine lures and neuter lures in group "fem.-neu.", t (14) = 2.85, $p$ = 0.01; while in group "masc.-neu.", the difference between masculine lures and neuter lures was not significant, t (14) = 0.46, $p$

= 0.65.

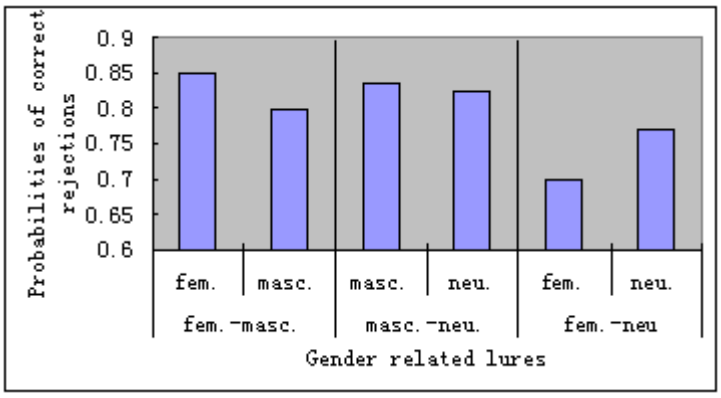



**Figure 6-2 Comparisons between gender-related lures within each gender combination group (Data across Times)**

In addition, the proportions of correct rejections to the lures were influenced significantly by number of repetitions, $F$ (2, 84) = 23.40, $MS$e = 0.44 (Table 6-3, details are in Appendix E - 8). A further contrast showed that proportions of correct rejections decreased along when the number of repetitions increased ($M_1$ = 0.85, $M_4$ = 0.79, $M_8$ = 0.74, $p$s ≤ 0.01). Similar to Experiment 2, different genders of lures did not cause different proportions of correct rejections, $F$ (2, 84) = 1.03, $MSe$ = 0.02.

Finally, the difference among different Gender combination groups (group "fem.-masc.", group "masc.-neu." and group "fem.-neu.") was marginally significant, $F$ (2, 42) = 3.03, $MSe$ = 0.51, $p$ = 0.06. Multiple comparisons showed that the proportion of correct rejection to group "masc.-neu." was significantly higher than that to group "fem.-neu." ($p$ = 0.02), but not significantly higher than that to group "fem.-masc.", at alpha level 0.05 ($p$ = 0.58).

### 6. 2. 2. False Alarms to Gender-Related and Gender-Unrelated Lures

There were three kinds of false alarms to the gender-related and gender-unrelated lures: responses of "Remember", "Know" and "Guess". Since the proportions of responses "Remember" and "Know" were too low to be further analyzed (data were mainly below 5%), further analysis focused on the proportion of response "Guess". Similar to the analysis described above, statistical analyses here included 3 (genders of lures: feminine, masculine, neuter) × 3 (gender combination groups: "fem. - masc.", "masc. - neu.", and "fem. - neu.") × 3 (Number of repetitions: one, four, eight times) mixed analyses of variance (ANOVAs).

First, the interaction between "gender of lures" and "gender combination groups" was not significant, $F$ (4, 84) = 1.55, $MSe$ = 0.02, $p$ = 0.20. In other words, grammatical gender effects were not found in "Guess" answers (Figure 6-3, details are in Appendix E - 8). Secondly, analyses of "Guess" responses showed a significant number-of-repetition effect, $F$ (2, 84) = 8.52, $MSe$ = 0.15. In contrast to the results from correct rejections, the proportion of correct rejections increased as the repetition of the relative learned word increased ($p_{1, 4} = 0.06$, $p_{4, 8} = 0.04$, $p_{1, 8} \leq 0.01$).

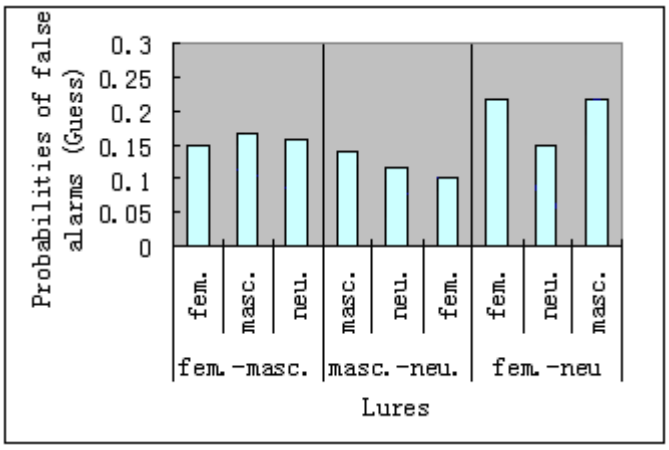



**Figure 6-3 Comparisons of false alarms among different lures within each gender combination group (Data across Times)**

### 6. 2. 3. Response Proportions of Targets (Hits)

As the analysis done to the proportions of targets in the previous Experiments, since the materials used in different gender combination groups (group "fem.-masc.", group "fem-neu.", and group "masc.-neu.") were not the same and cannot be combined together, all target responses belonging to different groups were analyzed separately. Within each gender combination group, statistical analyses included 2 (genders of targets: feminine and masculine in group "fem. - masc.", or masculine and neuter in group "masc. - neu.", or feminine and neuter in group "fem. - neu.") × 3 (number of repetitions: one, four, eight times) mixed analyses of variance (ANOVAs).

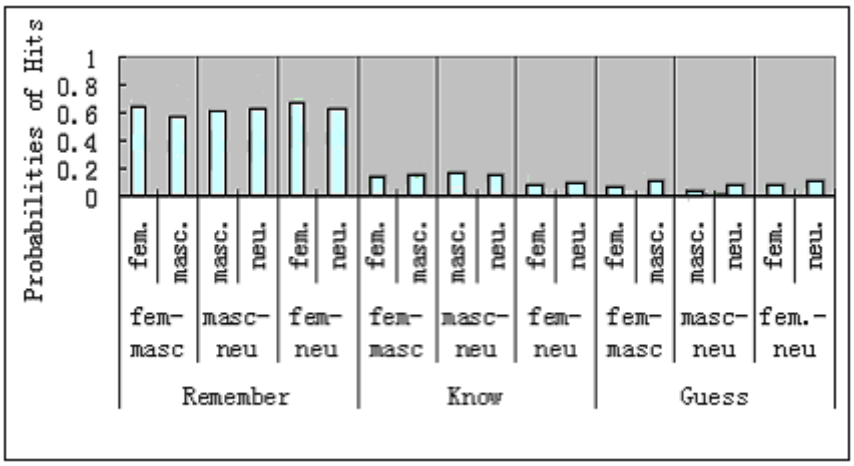



**Figure 6-4 Comparisons between targets within each gender combination group (Data across Times)**

Across different genders (feminine, masculine and neuter), an exact repetition effect was investigated. Results showed that more hits of "Remember" were found when the pseudo-determiner noun phrases were repeated eight times or four times rather than one time, $F$ (2, 178) = 127.13, $MSe$ = 4.37. In addition, hits of "Guess" were also influenced by the number of repetitions, $F$ (2, 178) = 10.09, $MSe$ = 0.15; by contrast, hits of "Know" did not vary with different number of repetitions, F < 1.

For detecting the difference between hits of noun phrases which belong to two different genders in the same group (Figure 6-4, details are in Appendix E - 9), repeated measures were applied to each target group. Targets belong to different genders in each group, for example, feminine targets and masculine targets in group "fem.-masc.", were analyzed. Across different number of repetitions, not only with hits of "Remember', but also with hits of "Know" and hits of "Guess", no significant difference was found between masculine and neuter in the group "masc. - neu.". In addition, no difference was found between feminine and masculine in the group "fem.

- masc.", or between feminine and neuter in the group "fem. - neu." ($F$s $\leq 1.72$[11]).

## 6. 3. Discussion

Firstly, as predicted, in this experiment, there were more correct rejections for gender-unrelated than related lures, when targets included masculine words and neuter words. However, no such effect occurred when targets included either feminine words and masculine words, or feminine words and neuter words. In other words, grammatical gender effects influences recognition tasks in the three-gender-condition only when masculine and neuter items were displayed in the learning stage. Such results indicated that, compared with masculine and neuter words, feminine words can be correctly rejected more frequently when it is the unrelated gender.

In addition, as gender-related lures, feminine words were more frequently rejected correctly than masculine words in group "fem. – masc.", but less frequently correctly rejected than neuter words in group "fem. – neu."; however, difference between the proportion of correct rejection to masculine and neuter words in group "masc.-neu." was not found.. Namely, it appears that the semantic category information but not the grammatical gender information was used as the recognition cue in the recognition task, when masculine and neuter items, which used the same pseudo determiner "elm", were displayed in the learning stage. Therefore, neither masculine nor neuter lures led to higher correct rejections. By contrast, besides the semantic category information, gender information can be also activated in the

[11] Dependent t-tests showed the same results.

recognition stage in group “fem. – masc.” and group “fem. – neu.”, but only one of the genders is better remembered than the other one so that more false alarms occurred. As mentioned in the previous experiments, different proportions of correct rejections between feminine and masculine lures and between feminine and neuter lures were called “gender conflicts” in this experiment. It should be noted that, different from Experiment 2, in which conflicts were found only when targets were repeated four times, in Experiment 4, such conflicts were found even across different number of repetitions. In short, using masculine and neuter targets apparently facilitated identifying gender-unrelated lures, while using either feminine and masculine or feminine and neuter targets did not. However, in this experiment, pseudo-determiners enhanced the gender conflicts between feminine and masculine items, and between feminine and neuter items. Even when words which belong to the same semantic category were repeated only once, gender conflict still occurred because gender information was activated by the different pseudo-determiners “elm”, and “elme”.

Further evidence for the important role played by pseudo-determiners came from the difference between masculine and neuter lures. Conflicts found in Experiment 3 between rejecting masculine lures and neuter lures in group “masc.-neu.” vanished in Experiment 4. Pseudo-determiners weakened the different gender information of masculine and neuter so that results consisted mainly with what has been demonstrated on false memory. In addition, arranging for the same determiners to apply to both masculine nouns and neuter nouns, resulted in an effect similar to what

was found with bare nouns (Experiment 2). As predicted, similar results were found with masculine and neuter items by using the same ending determiners (Experiment 4) or not (Experiment 2). Since masculine and neuter words use the same ending determiners more frequently in daily speaking, the gender information of these words cannot be identified very easily when they are displayed without the gender-specific determiners.

Secondly, a number-of-repetition effect was found not only with lures but also with targets. The more often the targets were repeated, the lower proportion of correct rejections to the gender-related lures occurred. On the contrary, the higher the number of repeated times, the more hits of "Remember" and of "Guess". Again, hits of "Know" were not so influenced by varying number of repetitions as hits of "Remember" and hits of "Guess".

Finally, although feminine words, as gender-related lures, were more likely to be rejected correctly than masculine words in the group "fem. - masc.", and less likely to be correctly rejected than neuter words in the group "fem. - neu.", the results of this experiment did not indicate any gender priority. None of the genders are consistently more memorable than the others.

# 7. Experiment 5

An eye-tracking system is a logical method to capture the participants' decision-making processes during the recognition task. Therefore, eye-tracking data was used to indicate how semantic category effects and grammatical gender effects are used by the participants in their recognition of the target words.

In this experiment, as the previous experiments, either feminine and masculine nouns, or masculine and neuter nouns, or feminine and neuter nouns were displayed in the learning stage. In the recognition stage, participants were instructed to differentiate the learned words from feminine, masculine, and neuter lures. Feminine and masculine lures are gender-related lures when feminine and masculine words have been displayed in the learning stage; masculine and neuter lures are gender-related lures when masculine and neuter words have been learned; and feminine and neuter lures are gender-related lures when feminine and neuter words have been presented.

## 7. 1. Method

**Participants:** Thirty[12] native German speakers from Heidelberg University (13 male, 17 female, mean age 22.13, *SD* = 2.37) participated in the experiment.

**Materials:** Eighteen semantic categories of six members each, including two feminine words, two masculine words and two neuter words, first arranged in Experiment 2, were again used.

[12] The sample size of 30 yields a statistical power of .9085, given an effect size of .25 and an alpha-level of .05 (analysis based on G*Power; see Faul et al., 2007).

**Procedure:** Study items were presented on Apple Macintosh computers. Each word remained on the screen for one second and was replaced after an interval of one second by the next item. Participants were instructed to read the words silently as they appeared on the screen and to bear in mind that they would later be given some form of memory test, the precise nature of which was not specified. Presentation of the 156 trials took approximately ten minutes. Participants were then provided with a distractor task for 15 minutes. Following this, they were given a recognition test.

In the recognition stage, 54 recognition pages, including 27 target pages and 27 lure pages, were presented by a program written in Python 2.0. On each target page, there was one target word (the learned word, e.g. *Katze* - cat) and two gender-related lures (e.g. *Woche* - week and *Fahne* - flag, which were not learned and were semantically unrelated to targets). On each lure page, all three words, one feminine word (e.g. *Ziege* - goat), one masculine word (e.g. *Hund* - dog) and one neuter word (e.g. *Pferd* - horse), were semantically related to the target word (e.g. *Katze* - cat). Participants were asked to select the word they had learned from each page by pressing the keyboard 1, 2, 3, or 4: If they believed that none of the three items on the page had been learned in the earlier phase, they should press "4". If they believed that an item did appear in the earlier task, they should press either 1 (the word on the first line), 2 (the word on the second line), or 3 (the word on the third line).


(A): Example of page with target (B): Example of page with only lures

**Figure 7-1 Examples of Materials Used in Recognition Task**

All of the pages in the recognition task were presented randomly. The current page disappeared when a keystroke was recorded, and the next page appeared 1000ms after that. All of the keystroke data and eye movement tracking were recorded by the software DataViewer and Eye-tracking Ⅱ.

## 7. 2. Results

### 7. 2. 1. Results of Keystrokes

The experimental factors for targets resulted in a 2 (gender conditions: gender-related vs. gender-unrelated lures) x 3 (number of repetitions: targets repeated one, four, or eight times) ANOVA test. Experiment factors to the lures resulted in a 2 (gender conditions: gender-related vs. gender-unrelated lures) x 3 (gender combination groups: words learned in the first stage were “fem.-masc.”, “fem.-neu.”, or “masc.-neu.”) x 3 (number of repetitions: targets repeated one, four, or eight times) mixed ANOVA. An alpha level of .05 was used in all statistical analyses. The proportions of correct and incorrect responses to targets and lures are displayed in Table 7-1, and Table 7-2.

**Table 7-1 Response Proportions (*p*) as a Function of Exact Repetition for Targets**

| | | Target group "fem.-masc." | | | Target group "fem.-neu." | | | Target group "masc.-neu." | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| F | Hits | 0.55 | 0.88 | 0.93 | 0.49 | 0.80 | 0.90 | / | / | / |
| | LA | 0.00 | 0.05 | 0.00 | 0.03 | 0.03 | 0.02 | / | / | / |
| | Miss | 0.45 | 0.07 | 0.07 | 0.48 | 0.17 | 0.08 | / | / | / |
| M | Hits | 0.72 | 0.92 | 0.95 | / | / | / | 0.78 | 0.94 | 0.97 |
| | LA | 0.00 | 0.03 | 0.02 | / | / | / | 0.02 | 0.03 | 0.00 |
| | Miss | 0.28 | 0.05 | 0.03 | / | / | / | 0.20 | 0.03 | 0.03 |
| N | Hits | / | / | / | 0.72 | 0.90 | 0.88 | 0.88 | 0.97 | 0.98 |
| | LA | / | / | / | 0.02 | 0.02 | 0.02 | 0.00 | 0.00 | 0.02 |
| | Miss | / | / | / | 0.26 | 0.08 | 0.10 | 0.12 | 0.03 | 0.00 |

Note: F: feminine targets, M: masculine targets, N: neuter targets.

LAs were also "misses" in this experiment. It means participants falsely identified lures that were gender related but semantically unrelated to the targets.

**Table 7-2 Response Proportions (*p*) as a Function of Exact Repetition for Lures**

| | | Lure group "fem.-masc." | | | Lure group "fem.-neu." | | | Lure group "masc.-neu." | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 4 | 8 | 1 | 4 | 8 | 1 | 4 | 8 |
| Correct rejections | | 0.94 | 0.90 | 0.85 | 0.88 | 0.88 | 0.83 | 0.92 | 0.86 | 0.90 |
| False alarms | GRL | 0.03 | 0.07 | 0.05 | 0.05 | 0.08 | 0.06 | 0.03 | 0.12 | 0.07 |
| | GUL | 0.03 | 0.03 | 0.1 | 0.07 | 0.04 | 0.1 | 0.05 | 0.02 | 0.03 |

Note: F: feminine lures, M: masculine lures, N: neuter lures.

GRL = grammatical-gender-related lures, GUL = grammatical-gender-unrelated lures. Both GRL and GUL were semantically related to targets.

Data retrieved from the lure pages were first analyzed. Words on these pages were only semantically related to targets, but were not targets themselves. In addition to being semantically related to targets, lures used in these lure pages were either gender-related or gender-unrelated to targets, according to what was displayed in the learning stage. The analysis of variance with the factors for lures showed no main effects of gender and number of repetitions, $F$s $\leq$ 1.31, and no interaction between genders (feminine, masculine, and neuter) and gender combination groups (group "fem.-masc.", group "fem.-neu.", and group "masc.-neu."), $F_{(4, 54)}$ =1.23, $MSe$ =0.01.

The analysis of hits to target pages showed a significant difference between two genders within the same gender combination group. That is, the proportion of hits to feminine targets were lower than those to masculine targets and to neuter targets, in group "fem.-masc.", $F_{(1, 9)}$ = 5.83, $MSe$ = 0.08, and in group "fem.-neu.", $F_{(1, 9)}$ =

10.57, *Mse* = 0.19, respectively; in group "masc.-neu.", neuter targets caused a higher proportion of hits, $F_{(1, 9)} = 5.65$, *Mse* = 0.05.

In addition, on each target page, there were two lures which had the same gender as the target word (gender-related lures). The target and the lure words were semantically unrelated. Thus, there were two kinds of misses: (1) rejecting both the target word and the gender-related lures; (2) rejecting the target word but accepting one of the lures. For investigating whether participants were more likely to falsely accept the gender-related lures than correctly reject them, these two kinds of misses were compared. However, the analysis of variance with the factors showed no significant difference between these two kinds of misses in the experimental conditions "fem.-masc.", group "fem.-neu.", or group "masc.-neu.", $F$s $\leq 2.09$.

The ANOVA revealed a main effect, the number of repetitions effect. The more often the targets were repeated in the learning phrase, the fewer misses occurred in the recognition stage. This number of repetitions effect was found in all grammatical gender conditions: "fem.-masc.", $F_{(1, 18)} = 28.56$, *Mse* = 0.56, "fem.-neu.", $F_{(1, 18)} = 22.23$, *Mse* = 0.47, "masc.-neu.", $F_{(1, 18)} = 8.14$, *Mse* =0.11.

### 7. 2. 2. Data of Eye-Tracking

#### 7. 2. 2. 1. Analysis of Pages that Included only Lures

There were eighteen lure pages. In each page, one feminine word, one masculine word and one neuter word were semantically related to one learned word displayed in the learning stage. In the following analysis, the data pertaining to dwell time and fixation

count are discussed. It can be assumed that longer fixation times indicate target words that are more difficult to recall or lures that are more difficult to reject. All data were analyzed by a 2 (gender conditions: gender-related vs. gender-unrelated lures) x 3 (gender combination groups: words learned in the first stage were "fem.-masc.", "fem.-neu.", or "masc.-neu.") x 3 (number of repetitions: targets repeated one, four, or eight times) mixed ANOVA.

The ANOVA showed no effects of the within-subjects factors on the dependent variables (either genders or number of repetitions). By contrast, a main effect of the between-subject factor "gender combination groups" was found (Table 7-3) on the dwell time across different interest areas, $F_{(2, 15)} = 17.19$, $MSe = 563994.68$, and on the number of fixations, $F_{(2, 15)} = 18.38$, $MSe = 12.73$. However, there was no significant difference among gender combinations ("fem.-masc.", "fem.-neu.", "masc.-neu."), concerning the durations of first fixations in each specific interest area, $F \leq 1$.

**Table 7-3**

**Results from the mixed ANOVAS Test for Semantic lures**

| | | IA_Dwell_Time | | | IA_First_Fiation_Time | | | IA_Fixation_Count | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | df | F | p. | df | F | p. | df | F | p. |
| Within Group Factors | Genders | 2 | 0.813 | 0.453 | 2 | 1.159 | 0.327 | 2 | 0.080 | 0.924 |
| | Genders * Groups | 4 | 0.158 | 0.958 | 4 | 0.506 | 0.732 | 4 | 0.102 | 0.981 |
| | Error (Genders) | 30 | | | 30 | | | 30 | | |
| | Number of repetitions | 2 | 0.074 | 0.929 | 2 | 0.308 | 0.737 | 2 | 0.204 | 0.817 |
| | Number of repetitions * Groups | 4 | 0.815 | 0.525 | 4 | 0.379 | 0.822 | 4 | 0.535 | 0.711 |
| | Error (Number of repetitions) | 30 | | | 30 | | | 30 | | |
| | Genders * Number of repetitions | 4 | 1.728 | 0.156 | 4 | 2.384 | 0.061 | 4 | 1.607 | 0.184 |
| | Genders * Number of repetitions * Groups | 8 | 0.666 | 0.719 | 8 | 0.651 | 0.731 | 8 | 0.418 | 0.906 |
| | Error (Genders * Number of repetitions) | 60 | | | 60 | | | 60 | | |
| Between Group Factors | Groups | 2 | 17.191 | 0.000* | 2 | 0.721 | 0.502 | 2 | 18.383 | 0.000* |

### 7. 2. 2. 2. Analysis of Pages that Included Targets

Target pages were the pages with one learned word (target) and two gender-related lures. All lures used in these pages were semantically unrelated to targets. Since materials used in different gender combination groups were not the same (feminine and masculine words were used in group "fem.-masc.", feminine and neuter words were used in group "fem-neu.", and masculine and neuter words were used in group "masc.-neu."), and cannot be analyzed together across different combination groups, all target pages belonged to each gender combination group were analyzed separately.

Data pertaining to "Interest Area Dwell Time" (dwell time) were analyzed first (Figure 7-2).

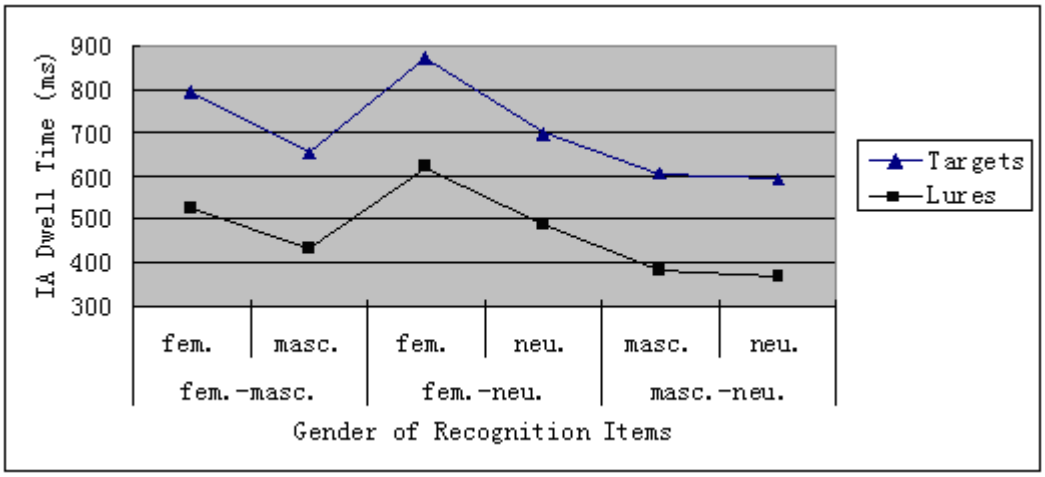



**Figure 7-2 Comparisons between feminine and masculine targets in group "fem.-masc.", feminine and neuter targets in group "fem.-neu.", masculine and neuter targets in group "masc.-neu." (IA Dwell Time)**

This data showed longer dwell time for feminine targets than either masculine targets (in group "fem.-masc."), $F_{(1,5)} = 17.38$, $MSe = 251684.55$, or neuter targets (in

group "fem.-neu."), $F_{(1,\ 5)}$ = 22.11, *MSe* = 412723.98. However, there was no difference between masculine and neuter words in group "masc.-neu.", $F \leq 1$. In addition, results of the analysis of the dwell time indicated longer fixations on the targets than on the lures. Targets were fixated longer in all gender combination groups: "fem.-masc.", $F_{(1,5)}$ = 16.47, *MSe* = 1085873.87, "fem.-neu.", $F_{(1,5)}$ = 206.03, *MSe* = 948810.52, "masc.-neu.", $F_{(1,5)}$ = 13.86, *MSe* = 909394.70.

Consistent with the results of previous studies, a main effect of number of repetitions was found. The more times the targets in the learning stage were repeated, the shorter the pages, including such targets, in the recognition phrase were fixated upon (i.e. in group "fem.-masc.", $F_{(2,\ 10)}$ = 4.10, *MSe* = 141073.13; in group "fem.-neu.", $F_{(2,\ 10)}$ = 13.22, *MSe* = 302445.81; in group "masc.-neu.", $F_{(2,\ 10)}$ = 6.46, *MSe* = 34187.70).

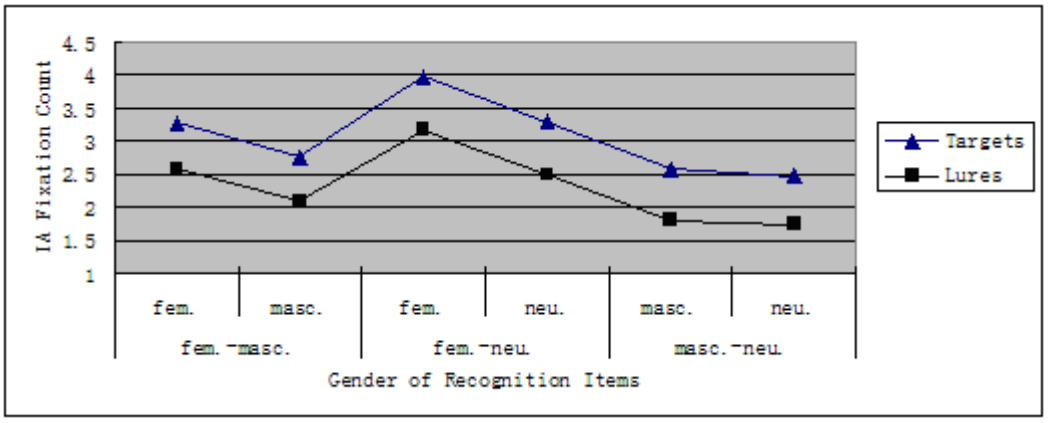



**Figure 7-3 Comparisons between two kinds of targets within each Gender combination group (IA Fixation Count)**

Moreover, similar main effects of gender and number of repetitions were achieved

by analyzing the numbers of fixations (Figure 7-3). That is, as targets, feminine words were more often fixated upon than either masculine words (in group "fem.-masc."), $F_{(1,5)} = 14.33$, $MSe = 4.68$, or neuter words (in group "fem.-neu."), $F_{(1,5)} = 38.23$, $MSe = 8.75$. Again, there was no difference between masculine and neuter words (in group "masc.-neu."), $F_{(1,5)} = 0.69$, $MSe = 0.11$. The number of fixations was influenced by the number of repetitions. The more times words were repeated in the learning stage, the fewer they were fixated upon in the recognition stage (i.e. in group "fem.-masc.", $F_{(2,10)} = 7.34$, $MSe = 3.03$; in group "fem.-neu.", $F_{(2,10)} = 16.97$, $MSe = 4.33$; in group "masc.-neu.", $F_{(2,10)} = 8.39$, $MSe = 0.43$). In addition, a significant difference between targets and lures was found in group "fem.-neu.", $F_{(1,5)} = 14.93$, $MSe = 11.68$; but not in either group "fem.-masc.", $F_{(1,5)} = 3.77$, $MSe = 8.59$, or group "masc.-neu.", $F_{(1,5)} = 5.39$, $MSe = 9.94$.

However, analyzing the first fixation duration (Figure 7-4), revealed no main effects of genders ($F$s $\leq 1.50$) and of the number of repetitions ($F$s $\leq 3.05$). There was, however, a significant difference between targets and lures: the first fixation time on the targets was again found to be longer than that on lures (i.e. in group "fem.-masc.", $F_{(1,5)} = 190.97$, $MSe = 67003.58$; in group "fem.-neu.", $F_{(1,5)} = 14.10$, $MSe = 7842.76$; in group "masc.-neu.", $F_{(1,5)} = 6.59$, $MSe = 44558.09$).

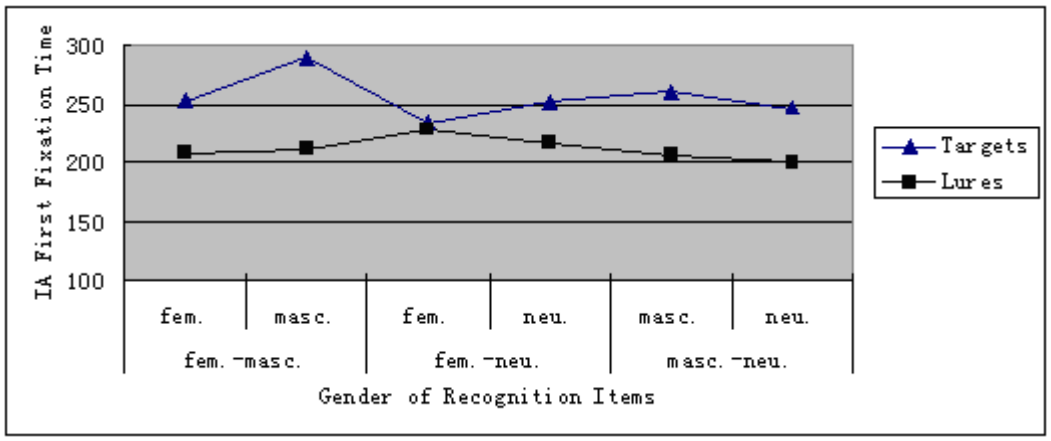



**Figure 7-4 Comparisons between two kinds of targets within each Gender combination group (IA First Fixation Time)**

## 7. 3. Discussion

Some of the results, which differed from what were achieved in the previous experiment, were considered to be caused by the different experimental design, and again supported the hypothesis that recognition tasks can be influenced by grammatical gender effects.

First of all, there was no interaction between gender of lures and gender combination of targets, neither in keystroke data nor by analyzing the eye-tracking data of lure-pages. This result is not consistent with the assumption of grammatical gender effects.

In this experiment, there were always two types of lures, gender-related lures and gender-unrelated lures, displayed in each recognition page (lure page). For example, two words that belong to the category four-foot animal [Katze / Cat (feminine), Elefant / elephant (masculine)] were learnt in the learning task, whereas two other words from this category were used in the recognition task (Ziege / Goat (feminine),

Hund / Dog (masculine), Pferd / Horse (neuter)). Results showed no significant difference between gender-related lures and gender-unrelated lures because of the low proportion of false alarms to the lures. It indicated that participants can compare these lures concerning features of semantic category or grammatical gender or both, so that false alarms were more often avoided

However, grammatical gender effects were still supported by the analysis of target pages. Significant differences were found between hits for feminine and masculine words in group "fem.-masc.", and for feminine and neuter words in group "fem.-neu.". As targets, feminine words caused fewer hits than masculine and neuter words in the recognition task. Moreover, the eye-tracking data showed longer fixation times for pages that included one feminine target and two feminine lures, which were semantically unrelated to the target, than for pages included masculine or neuter items. By contrast, masculine and neuter targets did not differ in terms of hits in group "masc.-neu.".

It is assumed that, also as mentioned in the previous experiments, recognition of feminine words differs from masculine and neuter words in the recognition task. In the previous four experiments, feminine lures were rejected more consistently than either masculine or neuter lures according to the proportion of correct rejection. In other words, feminine is more easily identified during the recognition task. In this experiment, a similar conclusion has been supported by the longer fixation and fewer hits on feminine targets. That is, in each page that included a target, there is one target and two lures. All of these three items are of the same grammatical gender. On the

pages that included masculine or neuter items, the recognition tasks were primarily influenced by semantic category effects, since either masculine or neuter was easily processed during the recognition tasks. As a result, shorter fixation time and more hits occurred to the targets. By contrast, on the pages with feminine targets, identification was simpler, because the recognition tasks were influenced not only by semantic category effects but also grammatical gender effects. Since the gender information (feminine) was processed during both the learned stages and the recognition task, feminine targets and lures were longer fixated and the gender-related lures were more often falsely identified as the targets.

Furthermore, the proportion of hits in this experiment was influenced by the number of repetitions. The more a target was repeated (hence, a higher number of repetitions), the more hits occurred. By contrast, by analyzing the eye-tracking data, the dwell time and the number of fixations decreased when the number of repetitions increased. These results are consistent with what was found in the preceding studies.

# 8. General Comparisons and General Discussion

## 8. 1. Comparisons among the Previous Experiments

Since the nouns used in experiment 2, 3, and 4 were the same, though with different determiner conditions (bare nouns, nouns with definite determiners, or nouns with pseudo-determiners), were the same, the results of lures from these experiments are ought to be compared against one another. Because of the low proportions of false alarms of the answers “Remember” and “Know” to the lures in experiment 4, only

false alarms of "Guess" and correct rejections were compared.

This analysis showed no significant effect based on the experiment design when the data of gender-related lures were analyzed, $F$ (2, 126) = 1.42, $MSe$ = 0.25 (Figure 8-1; details are in Appendix E - 10). The results of comparing gender-unrelated lures revealed that the proportion of correct rejections was influenced by different experiment designs, $F$ (2, 126) = 6.49, $MSe$ = 0.31. That is, the proportion of correct rejections to gender-unrelated lures in experiment 2 was higher than that in either experiment 3, $F$ (1, 84) = 8.71, $MSe$ = 0.38, or experiment 4, $F$ (1, 84) =12.00, $MSe$ = 0.55; by contrast, Experiment 3 and 4 had a similar proportions of correct rejections, $F$ (1, 84) = 0.28, $MSe$ = 0.02.

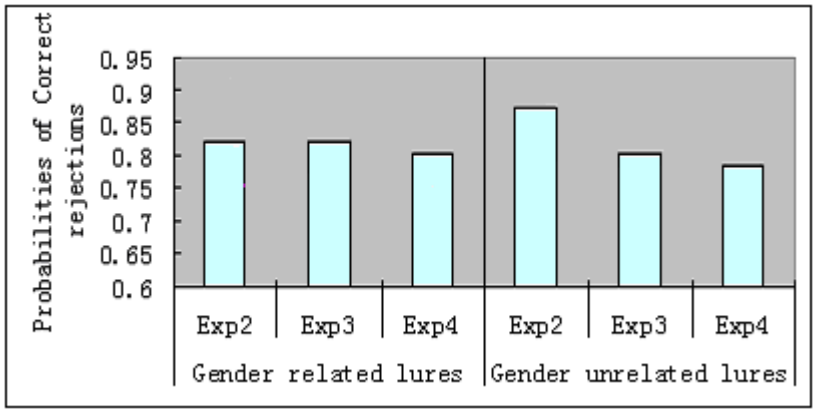



**Figure 8-1 Comparisons of Correct Rejections to different types of lures among Experiment 2, 3, and 4**

In contrast to the correct rejections, more "Guesses" were found in experiment 4 than in experiment 2; the difference among the designs of the different experiments were marginally significant, $F$ (2, 126) = 3.04, $MSe$ = 0.27, $p$ = 0.05. Results indicated a lower proportion of "Guesses" to gender-unrelated lures in experiment 2 than in

experiment 3, $F(1, 84) = 3.82$, $MSe = 0.12$, and experiment 4, $F(1, 84) = 13.83$, $MSe = 0.38$. In contrast to the results of correct rejections, the difference between experiment 3 and 4 was found to be significant, $F(1, 84) = 4.74$, $MSe = 0.16$ (Figure 8-2, details are in Appendix E - 11).

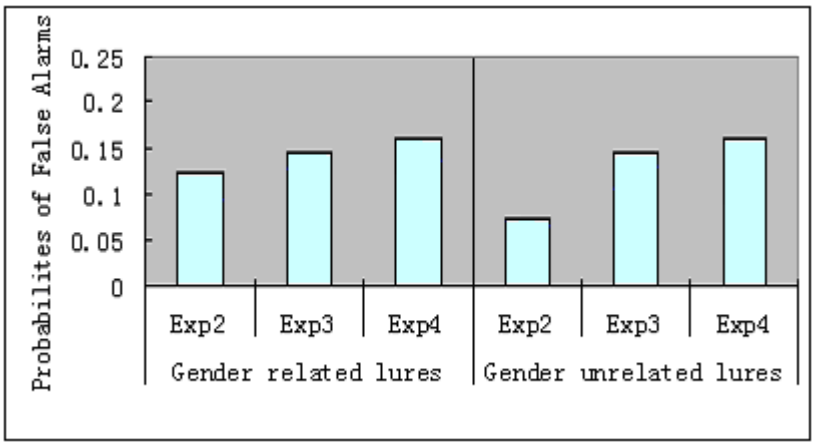



**Figure 8-2 Comparisons of False Alarms (Guess) to different types of lures among Experiment 2, 3, and 4**

It should be noted that, no matter which kind of answers were detected (correct rejections or "Guesses"), feminine lures were found to lead similar proportions in all three experiments (experiment 2, 3, and 4), when they acted as the gender-unrelated lures, $p$'s $\geq 0.19$. In contrast, the different experiment designs primarily affected masculine lures and neuter lures when they were gender-unrelated lures, $p$'s $\leq 0.03$.

## 8. 2. General Discussion

Recent research (e.g., Dewhurst et al., 1999; Dewhurst, 2001; Roediger & McDermott, 1995) has shown that participants in memory experiments incorrectly recognize non-

studied items if they are semantically related to items presented at the learning stage. The present studies investigated this false memory with grammatical gender: using German and the distinction among remembering, knowing and guessing to investigate the effect of grammatical gender in recognition memory.

In experiment 1, feminine lures were gender-related to the targets while masculine lures were gender-unrelated. Grammatical gender effects were detected with a two-gender condition. The most important result achieved was that feminine lures led to more false alarms than masculine lures, although both masculine and feminine lures were semantically related to targets.

As we know, phonological marking and semantic category are the two main factors that influence the lexical access. Sommers and Lewis (1999) stated that false memories may be caused by phonological as well as semantic similarity between target items and distractors. The finding from experiment 1 showed that, as one of the essential layers, which is as important as phonological marking and semantic category in producing nouns in the "Lemma model" (e.g. Roelofs, 1992; Levelt et al., 1999), grammatical gender can be one of the factors that influence the false memory. With two different genders in experiment 1, the more related factors between lures and targets (gender-related lures share two related factors with targets, semantic category and grammatical gender, while gender-unrelated lures share only one related factors with targets), the more false alarms occur to the lures. In addition, by analyzing the data of targets and lures across different grammatical genders, main effects are consistent with findings reported by Dewhurst and Anderson (1999).

Experiment 2, 3, 4 and 5 investigated the grammatical gender effects with a three-gender condition. Grammatical gender effects, which were expected, were not clear when the repetition of materials in the learning stage were massed, but was found in some specific situations when the repetition of materials were spaced. In addition, results from these three experiments indicated explicit effects from the mark of grammatical gender (determiners), different distinctive degrees among different genders, and recognition conflicts between genders in the learning stage.

Results achieved in experiment 2, 3, and 4 were compared. Grammatical gender effects were revealed in experiment 2 and 4 but not in experiment 3. As gender-unrelated lures, feminine nouns were more correctly rejected than masculine and neuter nouns, which were the gender-related lures. Moreover, such an effect was found in neither "fem.-masc." nor "fem.-neu." groups. Such results indicated that the feminine was more easily identified and used as a gender cue than either masculine or neuter in the recognition tasks. In the recognition task, once the gender is identified to be the non-studied, the words belonging to that gender is easily rejected. Therefore, as a gender-unrelated lure, a feminine word can be correctly rejected because of its gender cue, even if it is semantic related to the targets; by contrast, masculine and neuter words do not function as a gender cue as easily, so recognition tasks with masculine and neuter words are influenced mainly by the semantic category effects. Thus, more incorrect rejections of the target words occurred with masculine and neuter targets. Such conclusions are supported by the results from experiment 5. When the targets were identified from the other two same-gender lures, the

identifications were influenced either by grammatical gender effects (lures) or by both semantic category effects and grammatical gender effects (targets). Since feminine is more easily processed as the gender cue during the recognition tasks, feminine lures are more difficult to reject than either masculine or neuter lures. Therefore, pages with feminine words were fixated upon longer, according to the eye tracking data, and feminine targets correctly hit less often.

Moreover, results achieved in experiment 3 and 4 indicated that the grammatical gender effects which occurred during the recognition tasks might be a result of the effects of definite and indefinite determiners. As mentioned by Taft et al. (1998), representation of the article stored in mind plays a role in gender access. A similar influence was revealed in the recognition tasks. Namely, results demonstrated the relationship between articles and nouns in German. In German sentences, masculine and neuter have more opportunities to share the same determiners. For example, they share the same indefinite article "ein" and the same negative determiner "kein", when they are nominative in sentences; adjectives used by masculine and neuter nouns share the same ending "-em" when they are dative. From the indefinite article, the negative determiner or the ending of the adjective they use, gender of the nouns can not be identified; By contrast, feminine nouns use the indefinite article "eine" and the negative determiner "keine" when they are nominative in sentences. In addition, although the ending of adjectives used in the dative position is "-er", which is the same as the ending of adjectives used by masculine when they are nominative, the gender of feminine nouns can still be easily identified from the context. Thus,

compared with the distinction between masculine and neuter, feminine is more distinctive. This conclusion is supported by comparisons among different determiner conditions. Determiners were used both in experiment 3 and 4. In experiment 4, masculine and neuter shared the same pseudo-determiner "elm" and grammatical gender effects were found similar to what was achieved in experiment 2.

**Table 8-1 Comparisons between Proportions of Correct Rejections to Gender-related lures within the same Group**

| Groups | Lures | Experiment 2* | Experiment 3** | Experiment 4** |
|---|---|---|---|---|
| "fem.-masc." | fem.vs. masc. | $p_{\text{fem.}} > p_{\text{masc.}}$ | $p_{\text{fem.}} > p_{\text{masc.}}$ | $p_{\text{fem.}} > p_{\text{masc.}}$ |
| "masc.-neu." | masc.vs. neu. | Not significant | $p_{\text{masc.}} > p_{\text{neu.}}$ | Not significant |
| "fem. - neu." | fem. vs. neu. | $p_{\text{neu.}} > p_{\text{fem.}}$ | $p_{\text{neu.}} > p_{\text{fem.}}$ | $p_{\text{neu.}} > p_{\text{fem.}}$ |

*Note*: * Such a tendency occurred only with repeated 4-time targets.

** Such a tendency occurred with different number of repetitions.

However, such effects were not found in experiment 3. The results revealed that, in experiment 3, the definite determiners enhanced the gender information of both masculine and neuter words, so that advantages of feminine achieved in experiment 2 and 4 disappeared. Firstly, the determiners enhanced gender information of masculine and neuter. This effect is called "gender conflicts". This means that words belonging to one gender are always more easily identified in the recognition task than the other one when two different genders are learned in the same stage. By comparing proportions of correct rejections to gender-related lures within the same group, conflicts between gender-related lures were found except masculine vs. neuter in

group “masc.-neu.” in experiment 2 and 4 (Table 8-1). That is, feminine lures were correctly rejected more frequently than masculine lures in group “fem.-masc.”, but correctly rejected less frequently than neuter lures in group “fem.-neu.”. However, in experiment 3, the appropriate definite articles enhanced the gender information of both masculine and neuter, so that conflicts between these two genders occurred. More correct rejections occurred with masculine lures than with neuter lures. In short, the appropriate definite articles helped masculine and neuter become identified.

Secondly, clear-gender-mark noun phrases (experiment 3 and 4) were correctly rejected less often, either masculine or neuter, than bare nouns (experiment 2), when they are gender-unrelated lures; however, the proportion of correct rejections to feminine lures are similar in these three experiments. This indicates that both semantic and grammatical gender features are activated when identifying a noun as a learned word in the recognition tasks. When feminine lures, masculine lures, and neuter lures share the same semantic features, grammatical gender feature is the important information used to identify a learned word. On the one hand, feminine is so easily used as a gender cue that it can be identified without paying extra attention to the gender information carried by the appropriate determiners in the learning and recognition tasks. Therefore, whether feminine words have clear gender marks or not, they are more easily rejected when their gender information does not exist in the learning stage; on the other hand, masculine and neuter are difficult to use as gender cues without appropriate determiners. The results of correct rejections to the gender-unrelated lures were influenced by the definite determiners displayed with the words.

Thirdly, clear-gender-mark noun phrases led more false "Guesses" to lures than bare nouns (experiment 2) by comparing gender-unrelated lures; moreover, these differences occurred when gender-unrelated lures were either masculine or neuter, but not feminine.

In addition, across different grammatical genders, the results achieved in this study are consistent with results in the recent research. That is, the semantic category influences recognition performance by enhancing the recollection component of recognition memory, as measured by "Remembers" (Dewhurst & Anderson, 1999).

To summarize, the present findings show that recognition performance is influenced not only by semantic category, but also grammatical gender effects. Grammatical gender can affect the recognition memory. With the traditional experiment design (experiment 1, 2, 4), grammatical gender-related lures primarily caused more false alarms than gender-unrelated lures. However, once the semantic category information was enhanced, such a grammatical gender effect disappeared (experiment 3). In addition, with the more sensitive experimental procedure (experiment 5), feminine was revealed to be most easily identified and used as one of the memory cues in the recognition tasks; by contrast, except using the explicit gender cue (experiment 3), grammatical gender effects were not significant for masculine and neuter.

# References

Andonova, E., D'Amico, S., Devescovi, A., & Bates, E., (2004). Gender and lexical access in Bulgarian. *Perception & Psychophysics*, 66, 496-507.

Badecker, W., Miozzo, M., & Zanuttini R., (1995). The two-stage model of lexical retrieval: evidence from a case of anomia with selective preservation of grammatical gender. *Cognition*, 57. 193-216.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L., (1996). Gender priming in Italian. *Perception & Psychophysics*, 58, 992-1004.

Bates, E., Devescovi, A., Pizzamiglio, L., D'Amico, S., & Hernandez, A., (1995). Gender and lexical access in Italian. *Perception & Psychophysics*, 57, 847-862.

Bentrovato, S., Devescovi, A., D'Amico, S., Wicha, N., & Bates, E., (2003). The effect of grammatical gender and semantic context on lexical access in Italian using a timed word-naming paradigm. *Journal of Psycholinguistic Research*, 32, 417-430.

Bölte, J., & Connine, C. M., (2004). Grammatical gender in spoken word recognition in German. *Perception & Psychophysics*, 66, 1018-1032.

Caramazza, A., (1997). How many levels of processing are there in lexical access? *Cognition*

*Neuropsychology*, 14, 177-208.

Caramazza, A., & Miozzo, M., (1997). The relation between syntactic and phonological knowledge in lexical access: Evidence from the 'tip-of-the-tongue' phenomenon. *Cognition*, 64, 309-343.

Colé, P., Pynte, J., & Andriamamonjy, P., (2003). Effect of grammatical gender on visual word recognition: Evidence from lexical decision and eye movement experiments. *Perception & Psychophysics*, 65, 407-419.

Comrie, B., (1999). Grammatical Gender Systems: A linguist's assessment. *Journal of Psycholinguistic Research*, 28, 457-466.

Corrêa, L. M. S., de A. Almeida, D. A., & Porto, R. S., (2004). On the representation of Portuguese gender-inflected words in the mental lexicon. *Brain and Language*, 90, 63-73.

Cubelli, R., Lotto, L., Paolieri, D., Girelli, M., & Job, R., (2005). Grammatical gender is selected in bare noun production: Evidence from the picture-word interference paradigm. *Journal of Memory and Language*, 53, 42-59.

Dahan, D., Swingley, D., Tanenhaus, M. K., & Magnuson, J. S., (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language*, 42, 465-480.

Dewhurst, S. A.; & Anderson, S. J., (1999). Effects of exact and category repetition in true and false recognition memory. *Memory & Cognition*, 27, 664-673.

Dewhurst, S. A., (2001). Category repetition and false recognition: effects of instance frequency and category size. *Journal of Memory & Language*, 44, 153-167.

Dewhurst, S. A., & Farrand, P., (2004). Investigating the phenomenological characteristics of false recognition for categorized words. *European Journal of Cognitive Psychology*, 16, 403-416.

Dodd, M. D., & Macleod, C. M., (2004). False recognition without intentional learning. *Psychonomic Bulletin & Review*. 11, 137-142.

Engebretson, P. H., (1995). Category effects on the discriminability of spatial location. Dissertation Abstracts International: Section B: *The Sciences & Engineering*, 55, 5586.

Farella-Bisch, S.; (2003). Generation and category effects on item and source recall: age-related differences. Dissertation Abstracts International: Section B: *The Sciences & Engineering*, 63, 6114.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

Ferrand, L., (2001). Grammatical gender is also on the tip of French tongues. Current Psychology Letters: *Behaviour, Brain & Cognition*, 5, 7-20.

Fletcher, B., (1983). The role of category information in word identification: a parallel decision model. *Memory & Cognition*, 11, 237-250.

Friederici, A. D., & Jacobsen, T., (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, 28, Special issue: Processing of grammatical gender, Part 1, 467-484.

Gardiner, J. M., (1988). Functional aspects of recollective experience. *Memory & Cognition,* 16, 309–313.

Gollan, H. T., & Frost, R., (2001). Two routes to grammatical gender: evidence from Hebrew. *Journal of Psycholinguistic Research*, 30, 627-651.

Irmen, L., & Roβberg, N., (2004). Gender markedness of language: the impact of grammatical and nonlinguistic information on the mental representation of person information. *Journal of Language and Social Psychology*, 23, 272-307.

Jacobsen, T., (1999). Effects of grammatical gender on picture and word naming: evidence from German. *Journal of Psycholinguistic Research*, 28, Special issue: Processing of grammatical

gender, Part 1, 499-514.

Larochelle, S., & Pineau, H.; (1994). Determinants of response times in the semantic verification task. *Journal of Memory & Language*, 33, 796-823.

Lemhöfer, K., Schriefers, H., & Jescheniak, J. D., (2006). The processing of free and bound gender-marked morphemes in speech production: evidence from Dutch. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32, 437-442.

Levelt, W. J. M, & Roelofs, A., Meyer, A. S., (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22, 1-75.

Lew-Williams, C., & Fernald, A., (2007). Young children learning spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, 18, 193-198.

Macoir, J., & Béland, R., (2004). Knowing its gender without knowing its name: differential access to lexical information in a jargonaphasic patient. *Neurocase*, 10, 471–482.

McDermott, K. B., (1996). The persistence of false memories in list recall. *Journal of Memory and Language*, 35, 212–230.

McDermott, K. B. & Watson, J. M., (2001). The rise and fall of false recall: the impact of

presentation duration. *Journal of Memory and Language*, 45, 160-176.

Miozzo, M., & Caramazza, A., (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 25, 907-922.

Parducci, A., & Wedell, D. H.; (1986). The category effect with rating scales: number of categories, number of stimuli, and method of presentation. *Journal of Experimental Psychology: Human perception & performance*, 12, 496-516.

Perlak, D., & Jarema, G., (2003). The recognition of gender-marked nouns and verbs in Polish-speaking aphasic patients. *Cortex*, 39, 383-403.

Radeau, M., & van Berkum, Jos J. A., (1996). Gender decision. *Language and Cognitive Processes*, 11, 605-610.

Roediger, H. L., III, & McDermott, K. B., (1995). Creating false memories: remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition,* 21, 803–814.

Roelofs, A., (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107-142.

Sagi, A., (1984). Grammatical gender, symbolic meaning and gender concept: recall, classification and preference tests. *Psychology & Human Development*, 1, 1-9.

Schriefers, H., & Teruel, E., (2000). Grammatical gender in noun phrase production: the gender interference effect in German. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1368-1377.

Sun, P.., & Ling Z., (1998). Category effects in long-term recognition task. *Acta Psychologica Sinica*, 30, 254-261.

Taft, M., & Meunier, F., (1998). Lexical representation of gender: a quasiregular domain. *Journal of Psycholinguistic Research*, 27, 23-45.

Tulving, E., (1985). Memory and consciousness. *Canadian Psychologist,* 26, 1–12.

Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K., (2005). Grammatical gender effects on Cognition: Implications for language learning and language use. *Journal of Experimental Psychology: General*, 134, 501-520.

Vigliocco, G., & Franck, J., (1999). When sex and syntax go hand in hand: gender agreement in language production. *Journal of Memory and Language*, 40, 455-478.

Vigliocco, G., Vinson, D. P., Indefrey, P., Levelt, W. J. M., & Hellwig, F., (2004). Role of grammatical gender and semantics in german word production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30, 483-497.

Wang, S., & Li, H., (1992). An experimental study of category effect in short-term memory retrieval. *Psychological Science (China)*, 6, 1-7, 12.

Watson, J. M., Balota, D. A., & Roediger III, H. L., (2003). Creating false memories with hybrid lists of semantic and phonological associates: over-additive false memories produced by converging associative networks. *Journal of Memory and Language*, 49, 95–118.

Weber, A., & Paris, G., (2004). The origin of the linguistic gender effect in spoken-word recognition: evidence from non-native listening. *The 10th Annual Conference on Architectures and Mechanisms of Language Processing*. 16-18.

# Appendix A: Materials of Exp 1

**Appendix A. Materials of EXP 1**

| Category | Targets | F_Lures | M_Lures |
|---|---|---|---|
| Units of time | Woche | Stunde | Tag |
| 4-foot animals | Katze | Ziege | Hund |
| Vegetables | Kartoffel | Tomate | Salat |
| Musical instruments | Flöte | Violine | Flügel |
| Birds | Taube | Krähe | Adler |
| Trees | Eiche | Buche | Ahorn |
| Fruits | Ananas | Kirsche | Apfel |
| Insects | Biene | Ameise | Käfer |
| Flowers | Lilie | Rose | Jasmin |
| Kitchen utensils | Gabel | Tasse | Löffel |
| Clothing | Jacke | Jeans | Hut |
| Parts of the body | Schulter | Hand | Bauch |
| Materials | Seide | Wolle | Pelz |
| Tools | Axt | Schraube | Hammer |
| Furniture | Couch | Kommode | Stuhl |
| home appliances | Lampe | Kamera | Backofen |
| Weapons | Bombe | Pistole | Panzer |
| office supplies | Kreide | Akte | Locher |
| geometry | Kugel | Kurve | Würfel |
| algebra | Parabel | Achse | Parameter |
| smoking | Zigarre | Asche | Stumpen |
| Religion | Kapelle | Bibel | Altar |
| Studio | Palette | Statue | Pinsel |
| Book | Seite | Zeile | Untertitel |
| Innards | Lunge | Leber | Magen |
| Parts of eye | Linse | Iris | Sehnerv |
| Diseases | Erkältung | Grippe | Krebs |
| Post | Adresse | Telegrafie | Briefkasten |
| Sweetmeat | Schokolade | Torte | Pudding |
| Drinks | Limonade | Cola | Whisky |
| Physics | Verlagerung | Brechung | Druck |
| Negative Emotion | Wut | Furcht | Zorn |
| Positive Emotions | Liebe | Freude | Stolz |
| Bad weather | Flut | Lawine | Hagel |
| Electricity | Elektrizität | Schaltung | Magnetismus |
| Chemistry | Oxidation | Destillation | Phosphor |

# Appendix B: Results of the Pretest and the Main Study of Experiment 2

## Appendix B - 1 Materials of EXP 2

| F_Targets | F_Lures | M_Target | M_Lures | N_Target | N_Lures |
|---|---|---|---|---|---|
| Katze | Ziege | Hund | Elefant | Lamm | Pferd |
| Flöte | Violine | Flügel | Kontrabaß | Klavier | Saxophon |
| Lilie | Rose | Jasmin | Flieder | Violett | Gänseblümchen |
| Schulter | Hand | Bauch | Hals | Gehirn | Ohr |
| Seide | Wolle | Pelz | Samt | Leinen | Nylon |
| Kugel | Kurve | Würfel | Bogen | Viereck | Prisma |
| Palette | Statue | Pinsel | Zeichenblock | Plakat | Porträt |
| Adresse | Briefmarke | Briefkasten | Postamt | Porto | Postfach |
| Limonade | Cola | Whisky | Tee | Bier | Mineralwasser |
| Oxidation | Destillation | Phosphor | Stickstoff | Platin | Quecksilber |
| Prosa | Fabel | Roman | Mythos | Märchen | Gedicht |
| Straßenbahn | Fähre | Wagen | Zug | Schiff | Boot |
| Wohnung | Herberge | Gasthof | Raum | Haus | Hotel |
| Terrasse | Küche | Balkon | Keller | Zimmer | Bad |
| Couch | Kommode | Stuhl | Teppich | Regal | Sofa |
| Bombe | Pistole | Panzer | Revolver | Geschoß | Geschütz |
| Jacke | Jeans | Hut | Schal | Halstuch | Kleid |
| Gabel | Tasse | Löffel | Teller | Messer | Tablett |

## Appendix B - 3 Means of the Main Study in Experiment 2

### Appendix B - 3 - 1 Means of Targets in Group "fem.-masc."

| | Feminine Targets | | | | | | Masculine targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | | | | | |
| Remember | 0.30 | 0.30 | 0.74 | 0.39 | 0.82 | 0.19 | 0.30 | 0.16 | 0.67 | 0.29 | 0.76 | 0.23 |
| Know | 0.22 | 0.26 | 0.22 | 0.32 | 0.11 | 0.12 | 0.20 | 0.20 | 0.15 | 0.19 | 0.07 | 0.12 |
| Guess | 0.07 | 0.12 | 0.00 | 0.00 | 0.07 | 0.12 | 0.11 | 0.14 | 0.13 | 0.21 | 0.04 | 0.11 |
| No | 0.41 | 0.24 | 0.04 | 0.11 | 0.00 | 0.00 | 0.39 | 0.28 | 0.05 | 0.08 | 0.13 | 0.16 |
| Massed repetition | | | | | | | | | | | | |
| Remember | 0.38 | 0.21 | 0.54 | 0.17 | 0.46 | 0.20 | 0.29 | 0.23 | 0.40 | 0.29 | 0.56 | 0.28 |
| Know | 0.19 | 0.23 | 0.15 | 0.17 | 0.25 | 0.25 | 0.27 | 0.31 | 0.25 | 0.25 | 0.19 | 0.19 |
| Guess | 0.12 | 0.15 | 0.06 | 0.09 | 0.10 | 0.12 | 0.17 | 0.15 | 0.13 | 0.15 | 0.10 | 0.12 |
| No | 0.31 | 0.17 | 0.25 | 0.20 | 0.19 | 0.19 | 0.27 | 0.23 | 0.22 | 0.15 | 0.15 | 0.14 |

## Appendix B - 3 - 2 Means of Gender Related Lures in Group "fem.-masc."

| | Feminine Lures | | | | | | Masculine Lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | | | | | |
| Remember | 0.00 | 0.02 | 0.01 | 0.02 | 0.00 | 0.02 | 0.02 | 0.02 | 0.00 | 0.02 | 0.01 | 0.03 |
| Know | 0.00 | 0.02 | 0.01 | 0.03 | 0.01 | 0.03 | 0.00 | 0.02 | 0.00 | 0.02 | 0.01 | 0.02 |
| Guess | 0.05 | 0.04 | 0.11 | 0.04 | 0.09 | 0.03 | 0.06 | 0.04 | 0.02 | 0.04 | 0.13 | 0.04 |
| No | 0.95 | 0.05 | 0.87 | 0.05 | 0.90 | 0.05 | 0.92 | 0.04 | 0.98 | 0.05 | 0.85 | 0.05 |
| Massed repetition | | | | | | | | | | | | |
| Remember | 0.04 | 0.02 | 0.00 | 0.02 | 0.00 | 0.02 | 0.00 | 0.02 | 0.02 | 0.02 | 0.05 | 0.03 |
| Know | 0.00 | 0.02 | 0.11 | 0.03 | 0.07 | 0.02 | 0.07 | 0.02 | 0.09 | 0.02 | 0.07 | 0.02 |
| Guess | 0.13 | 0.04 | 0.06 | 0.04 | 0.15 | 0.04 | 0.09 | 0.04 | 0.07 | 0.04 | 0.14 | 0.04 |
| No | 0.83 | 0.05 | 0.83 | 0.05 | 0.78 | 0.05 | 0.84 | 0.04 | 0.82 | 0.05 | 0.74 | 0.05 |

## Appendix B - 3 - 3 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "fem.-masc."

| | Neuter lures | | | | | | Semantic unrelated lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | |
| Remember | 0.00 | 0.02 | 0.01 | 0.02 | 0.03 | 0.02 | 0.02 | 0.04 |
| Know | 0.00 | 0.02 | 0.09 | 0.03 | 0.02 | 0.02 | 0.02 | 0.03 |
| Guess | 0.07 | 0.05 | 0.05 | 0.04 | 0.02 | 0.04 | 0.04 | 0.05 |
| No | 0.93 | 0.06 | 0.85 | 0.05 | 0.93 | 0.05 | 0.92 | 0.06 |
| Massed repetition | | | | | | | | |
| Remember | 0.05 | 0.02 | 0.08 | 0.02 | 0.02 | 0.02 | 0.03 | 0.03 |
| Know | 0.02 | 0.02 | 0.08 | 0.03 | 0.05 | 0.02 | 0.08 | 0.07 |
| Guess | 0.19 | 0.05 | 0.11 | 0.04 | 0.06 | 0.04 | 0.10 | 0.12 |
| No | 0.74 | 0.06 | 0.73 | 0.05 | 0.87 | 0.05 | 0.79 | 0.20 |

## Appendix B - 3 - 4 Means of Targets in Group "Masc.-neu."

| | Masculine targets | | | | | | Neuter targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | | | | | |
| Remember | 0.43 | 0.25 | 0.70 | 0.25 | 0.80 | 0.20 | 0.41 | 0.31 | 0.78 | 0.20 | 0.74 | 0.27 |
| Know | 0.15 | 0.15 | 0.07 | 0.12 | 0.06 | 0.12 | 0.16 | 0.24 | 0.06 | 0.12 | 0.11 | 0.14 |
| Guess | 0.09 | 0.17 | 0.06 | 0.08 | 0.09 | 0.15 | 0.13 | 0.14 | 0.20 | 0.06 | 0.06 | 0.12 |
| No | 0.33 | 0.14 | 0.17 | 0.17 | 0.05 | 0.08 | 0.30 | 0.25 | 0.16 | 0.13 | 0.09 | 0.12 |
| Massed repetition | | | | | | | | | | | | |
| Remember | 0.63 | 0.26 | 0.50 | 0.18 | 0.73 | 0.27 | 0.65 | 0.33 | 0.66 | 0.22 | 0.69 | 0.21 |
| Know | 0.10 | 0.12 | 0.19 | 017 | 0.15 | 0.19 | 0.10 | 0.23 | 0.15 | 0.14 | 0.04 | 0.12 |
| Guess | 0.10 | 0.09 | 0.14 | 0.21 | 0.04 | 0.08 | 0.06 | 0.09 | 0.08 | 0.09 | 0.17 | 0.15 |
| No | 0.17 | 0.18 | 0.17 | 0.13 | 0.08 | 0.18 | 0.19 | 0.16 | 0.11 | 0.20 | 0.10 | 0.12 |

## Appendix B - 3 - 5 Means of Gender Related Lures in Group "Masc.-neu."

| | Masculine lures | | | | | | Neuter lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | | | | | |
| Remember | 0.01 | 0.02 | 0.02 | 0.02 | 0.08 | 0.03 | 0.07 | 0.02 | 0.07 | 0.02 | 0.03 | 0.00 |
| Know | 0.07 | 0.02 | 0.05 | 0.02 | 0.08 | 0.02 | 0.05 | 0.02 | 0.06 | 0.03 | 0.05 | 0.02 |
| Guess | 0.18 | 0.04 | 0.15 | 0.04 | 0.21 | 0.04 | 0.09 | 0.05 | 0.17 | 0.04 | 0.17 | 0.04 |
| No | 0.74 | 0.04 | 0.78 | 0.05 | 0.63 | 0.05 | 0.79 | 0.06 | 0.70 | 0.05 | 0.75 | 0.05 |
| Massed repetition | | | | | | | | | | | | |
| Remember | 0.02 | 0.02 | 0.02 | 0.02 | 0.05 | 0.03 | 0.05 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 |
| Know | 0.02 | 0.02 | 0.02 | 0.02 | 0.05 | 0.02 | 0.09 | 0.02 | 0.11 | 0.02 | 0.04 | 0.02 |
| Guess | 0.20 | 0.04 | 0.22 | 0.04 | 0.19 | 0.04 | 0.18 | 0.05 | 0.09 | 0.04 | 0.11 | 0.04 |
| No | 0.76 | 0.04 | 0.74 | 0.05 | 0.71 | 0.05 | 0.68 | 0.06 | 0.78 | 0.05 | 0.84 | 0.05 |

## Appendix B - 3 - 6 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "Masc.-neu."

| | Feminine lures | | | | | | Semantic unrelated lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | |
| Remember | 0.05 | 0.02 | 0.02 | 0.02 | 0.05 | 0.01 | 0.02 | 0.06 |
| Know | 0.00 | 0.02 | 0.02 | 0.03 | 0.07 | 0.03 | 0.03 | 0.03 |
| Guess | 0.08 | 0.04 | 0.11 | 0.04 | 0.11 | 0.03 | 0.09 | 0.10 |
| No | 0.87 | 0.05 | 0.86 | 0.05 | 0.77 | 0.05 | 0.86 | 0.10 |
| Massed repetition | | | | | | | | |
| Remember | 0.02 | 0.02 | 0.00 | 0.02 | 0.02 | 0.03 | 0.00 | 0.00 |
| Know | 0.07 | 0.02 | 0.09 | 0.03 | 0.05 | 0.03 | 0.05 | 0.08 |
| Guess | 0.14 | 0.04 | 0.09 | 0.04 | 0.16 | 0.03 | 0.15 | 0.14 |
| No | 0.77 | 0.05 | 0.82 | 0.05 | 0.77 | 0.05 | 0.80 | 0.21 |

## Appendix B - 3 - 7 Means of Targets in Group "fem.-neu."

| | Feminine targets | | | | | | Neuter targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | | | | | |
| Remember | 0.41 | 0.30 | 0.74 | 0.19 | 0.87 | 0.12 | 0.39 | 0.32 | 0.67 | 0.35 | 0.72 | 0.38 |
| Know | 0.17 | 0.29 | 0.13 | 0.11 | 0.04 | 0.11 | 0.20 | 0.32 | 0.07 | 0.12 | 0.06 | 0.08 |
| Guess | 0.11 | 0.12 | 0.04 | 0.07 | 0.02 | 0.06 | 0.04 | 0.11 | 0.07 | 0.17 | 0.07 | 0.15 |
| No | 0.31 | 0.29 | 0.09 | 0.17 | 0.07 | 0.17 | 0.37 | 0.32 | 0.19 | 0.21 | 0.15 | 0.29 |
| Massed repetition | | | | | | | | | | | | |
| Remember | 0.36 | 0.14 | 0.42 | 0.22 | 0.63 | 0.17 | 0.46 | 0.12 | 0.71 | 0.26 | 0.73 | 0.17 |
| Know | 0.10 | 0.12 | 0.04 | 0.08 | 0.06 | 0.09 | 0.08 | 0.13 | 0.08 | 0.13 | 0.04 | 0.08 |
| Guess | 0.21 | 0.31 | 0.10 | 0.15 | 0.10 | 0.15 | 0.13 | 0.11 | 0.08 | 0.13 | 0.19 | 0.17 |
| No | 0.33 | 0.26 | 0.44 | 0.20 | 0.21 | 0.15 | 0.33 | 0.20 | 0.13 | 0.19 | 0.04 | 0.08 |

**Appendix B - 3 - 8 Means of Gender Related Lures in Group "fem.-neu."**

| | Feminine lures | | | | | | Neuter lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | | | | | |
| Remember | 0.01 | 0.02 | 0.02 | 0.02 | 0.02 | 0.02 | 0.01 | 0.02 | 0.01 | 0.02 | 0.02 | 0.02 |
| Know | 0.00 | 0.02 | 0.01 | 0.03 | 0.02 | 0.03 | 0.03 | 0.02 | 0.02 | 0.03 | 0.08 | 0.02 |
| Guess | 0.09 | 0.04 | 0.12 | 0.04 | 0.19 | 0.04 | 0.08 | 0.05 | 0.21 | 0.04 | 0.11 | 0.04 |
| No | 0.90 | 0.05 | 0.85 | 0.05 | 0.77 | 0.05 | 0.88 | 0.06 | 0.76 | 0.05 | 0.79 | 0.05 |
| Massed repetition | | | | | | | | | | | | |
| Remember | 0.01 | 0.02 | 0.20 | 0.02 | 0.34 | 0.02 | 0.15 | 0.02 | 0.28 | 0.02 | 0.34 | 0.02 |
| Know | 0.00 | 0.02 | 0.11 | 0.03 | 0.11 | 0.03 | 0.02 | 0.02 | 0.05 | 0.03 | 0.05 | 0.02 |
| Guess | 0.13 | 0.04 | 0.07 | 0.04 | 0.07 | 0.03 | 0.12 | 0.05 | 0.12 | 0.04 | 0.07 | 0.04 |
| No | 0.86 | 0.05 | 0.62 | 0.05 | 0.48 | 0.05 | 0.71 | 0.06 | 0.55 | 0.05 | 0.54 | 0.05 |

**Appendix B - 3 - 9 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "fem.-neu."**

| | Masculine lures | | | | | | Semantic unrelated lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Spaced repetition | | | | | | | | |
| Remember | 0.01 | 0.02 | 0.01 | 0.02 | 0.06 | 0.03 | 0.06 | 0.05 |
| Know | 0.02 | 0.02 | 0.00 | 0.02 | 0.00 | 0.02 | 0.04 | 0.04 |
| Guess | 0.06 | 0.04 | 0.06 | 0.04 | 0.10 | 0.04 | 0.09 | 0.10 |
| No | 0.91 | 0.04 | 0.93 | 0.05 | 0.84 | 0.05 | 0.81 | 0.14 |
| Massed repetition | | | | | | | | |
| Remember | 0.23 | 0.02 | 0.43 | 0.02 | 0.00 | 0.03 | 0.21 | 0.06 |
| Know | 0.13 | 0.02 | 0.13 | 0.02 | 0.01 | 0.02 | 0.06 | 0.05 |
| Guess | 0.06 | 0.04 | 0.13 | 0.04 | 0.15 | 0.04 | 0.17 | 0.15 |
| No | 0.58 | 0.04 | 0.31 | 0.05 | 0.84 | 0.05 | 0.56 | 0.16 |

## Appendix B - 4 Results of Lures

### Appendix B - 4 – 1 Test of Within-Subjects Effects for Semantic Lures (Answers Remember)

| | Sum of Squares | *df* | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Lures * Times | 0.162 | 4 | 0.040 | 9.954 | 0.000 |
| Lures*Times*Condition*Target | 0.804 | 8 | 0.101 | 24.741 | 0.000 |
| Lures*Times*Condition | 0.314 | 4 | 0.079 | 19.340 | 0.000 |
| Lures*Times*Target | 0.724 | 8 | 0.091 | 22.276 | 0.000 |
| Error (Lures*Times) | 1.365 | 336 | 0.004 | | |
| Lures | 0.113 | 2 | 0.057 | 9.325 | 0.000 |
| Lures*Condition*Target | 0.226 | 4 | 0.057 | 9.336 | 0.000 |
| Lures*Condition | 0.106 | 2 | 0.053 | 8.755 | 0.000 |
| Lures*Target | 0.261 | 4 | 0.065 | 10.741 | 0.000 |
| Error (Lures) | 1.019 | 168 | 0.006 | | |
| Times | 0.068 | 2 | 0.034 | 8.127 | 0.000 |
| Times*Condition*Target | 0.044 | 4 | 0.011 | 2.640 | 0.036 |
| Times*Condition | 0.035 | 2 | 0.018 | 4.251 | 0.016 |
| Times*Target | 0.012 | 4 | 0.003 | 0.700 | 0.593 |
| Error (Times) | 0.701 | 168 | 0.004 | | |

### Appendix B - 4 – 2 Test of Within-Subjects Effects for Semantic Lures (Answers Know)

| | Sum of Squares | *df* | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Lures * Times | 0.103 | 4 | 0.026 | 3.551 | 0.007 |
| Lures*Times*Condition*Target | 0.096 | 8 | 0.012 | 1.558 | 0.105 |
| Lures*Times*Condition | 0.030 | 4 | 0.007 | 1.036 | 0.388 |
| Lures*Times*Target | 0.151 | 8 | 0.019 | 2.621 | 0.009 |
| Error (Lures*Times) | 2.427 | 336 | 0.007 | | |
| Lures | 0.084 | 2 | 0.042 | 7.672 | 0.001 |
| Lures*Condition*Target | 0.025 | 4 | 0.006 | 1.166 | 0.328 |
| Lures*Condition | 0.063 | 2 | 0.032 | 5.823 | 0.004 |
| Lures*Target | o.o38 | 4 | 0.009 | 1.728 | 0.146 |
| Error (Lures) | 0.916 | 168 | 0.005 | | |
| Times | 0.010 | 2 | 0.005 | 0.750 | 0.474 |
| Times*Condition*Target | 0.159 | 4 | 0.040 | 6.199 | 0.000 |
| Times*Condition | 0.053 | 2 | 0.026 | 4.104 | 0.018 |
| Times*Target | 0.016 | 4 | 0.004 | 0.637 | 0.637 |
| Error (Times) | 1.079 | 168 | 0.006 | | |

**Appendix B - 4 - 3 Test of Within-Subjects Effects for Semantic Lures (Answers Guess)**

| | Sum of Squares | *df* | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Lures * Times | 0.209 | 4 | 0.052 | 3.364 | 0.010 |
| Lures*Times*Condition*Target | 0.128 | 8 | 0.016 | 1.037 | 0.408 |
| Lures*Times*Condition | 0.198 | 4 | 0.050 | 3.198 | 0.013 |
| Lures*Times*Target | 0.152 | 8 | 0.019 | 1.299 | 0.281 |
| Error (Lures*Times) | 5.206 | 336 | 0.015 | | |
| Lures | 0.032 | 2 | 0.016 | 1.207 | 0.302 |
| Lures*Condition*Target | 0.006 | 4 | 0.016 | 0.123 | 0.974 |
| Lures*Condition | 0.161 | 2 | 0.081 | 6.153 | 0.003 |
| Lures*Target | 0.081 | 4 | 0.020 | 1.551 | 0.190 |
| Error (Lures) | 2.203 | 168 | 0.013 | | |
| Times | 0.029 | 2 | 0.015 | 1.303 | 0.275 |
| Times*Condition*Target | 0.122 | 4 | 0.031 | 2.743 | 0.030 |
| Times*Condition | 0.031 | 2 | 0.015 | 1.376 | 0.255 |
| Times*Target | 0.319 | 4 | 0.080 | 7.153 | 0.000 |
| Error (Times) | 1.873 | 168 | 0.011 | | |

## Appendix B - 5 Results of Targets

### Appendix B - 5 - 1 Test of Within-Subjects Effects of Targets (Answers Remember)

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Times | 8.553 | 2 | 4.277 | 146.423 | 0.000 |
| Times*Condition*Target*Recognition | 0.685 | 12 | 0.057 | 1.953 | 0.028 |
| Times*Condition | 1.834 | 2 | 0.917 | 31.400 | 0.000 |
| Times*Target | 0.337 | 4 | 0.084 | 2.881 | 0.023 |
| Times*Recognition | 0.464 | 4 | 0.116 | 3.969 | 0.004 |
| Error (Times) | 9.813 | 336 | 0.029 | | |

Note: “Times” was the repetition times of targets, including 1, 4 and 8 times;

“Condition” was the repetition condition of targets, including spaced condition or massed condition;

“Target” was the genders of the displayed target words, including “feminine and masculine”, “feminine and neuter”, or “masculine and neuter”;

### Appendix B - 5 - 2 Test of Between-Subjects Effects of Targets (Answers Remember)

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Intercept | 181.633 | 1 | 181.633 | 1487.954 | 0.000 |
| Condition*Target*Recognition | 1.591 | 6 | 0.265 | 2.173 | 0.048 |
| Condition | 0.934 | 1 | 0.934 | 7.653 | 0.006 |
| Target | 0.584 | 2 | 0.292 | 2.391 | 0.095 |
| Recognition | 0.166 | 2 | 0.083 | 0.679 | 0.508 |
| Error | 20.508 | 168 | 0.122 | | |

**Appendix B - 5 - 3 Test of Within-Subjects Effects of Targets (Answers Know)**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Times | 0.449 | 2 | 0.224 | 10.832 | 0.000 |
| Times*Condition*Target*Recognition | 0.385 | 12 | 0.032 | 1.549 | 0.105 |
| Times*Condition | 0.236 | 2 | 0.118 | 5.689 | 0.004 |
| Times*Target | 0.055 | 4 | 0.014 | 0.661 | 0.620 |
| Times*Recognition | 0.057 | 4 | 0.014 | 0.685 | 0.603 |
| Error (Times) | 6.962 | 336 | 0.021 | | |

**Appendix B - 5 - 4 Test of Between-Subjects Effects of Targets (Answers Know)**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Intercept | 10.160 | 1 | 10.160 | 175.084 | 0.000 |
| Condition*Target*Recognition | 0.451 | 6 | 0.075 | 1.295 | 0.262 |
| Condition | 0.005 | 1 | 0.005 | 0.083 | 0.774 |
| Target | 0.777 | 2 | 0.388 | 6.695 | 0.002 |
| Recognition | 0.006 | 2 | 0.002 | 0.051 | 0.950 |
| Error | 9.749 | 168 | 0.058 | | |

**Appendix B - 5 - 5 Test of Within-Subjects Effects of Targets (Answers Guess)**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Times | 0.119 | 2 | 0.060 | 5.129 | 0.006 |
| Times*Condition*Target*Recognition | 0.366 | 12 | 0.031 | 2.628 | 0.002 |
| Times*Condition | 0.014 | 2 | 0.007 | 0.608 | 0.545 |
| Times*Target | 0.031 | 4 | 0.008 | 0.666 | 0.616 |
| Times*Recognition | 0.257 | 4 | 0.064 | 5.533 | 0.000 |
| Error (Times) | 3.901 | 366 | 0.012 | | |

**Appendix B - 5 - 6 Test of Between-Subjects Effects of Targets (Answers Guess)**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Intercept | 4.611 | 1 | 4.611 | 154.897 | 0.000 |
| Condition*Target*Recognition | 0.075 | 6 | 0.013 | 0.422 | 0.863 |
| Condition | 0.383 | 1 | 0.383 | 12.864 | 0.000 |
| Target | 0.041 | 2 | 0.021 | 0.689 | 0.503 |
| Recognition | 0.057 | 2 | 0.028 | 0.956 | 0.387 |
| Error | 5.001 | 168 | 0.030 | | |

# Appendix C: Results of Experiment 3

## Appendix C - 1 Means of the Main Study in Experiment 3

### Appendix C - 1 - 1 Means of Targets in Group "fem.-masc."

| Spaced repetition | Feminine Targets | | | | | | Masculine targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.47 | 0.35 | 0.86 | 0.19 | 0.86 | 0.23 | 0.36 | 0.24 | 0.67 | 0.33 | 0.85 | 0.30 |
| Know | 0.07 | 0.12 | 0.11 | 0.20 | 0.04 | 0.12 | 0.07 | 0.17 | 0.09 | 0.15 | 0.13 | 0.25 |
| Guess | 0.18 | 0.12 | 0.00 | 0.00 | 0.10 | 0.12 | 0.16 | 0.15 | 0.17 | 0.16 | 0.02 | 0.06 |
| No | 0.29 | 0.23 | 0.03 | 0.07 | 0.00 | 0.00 | 0.41 | 0.28 | 0.07 | 0.14 | 0.00 | 0.00 |

### Appendix C - 1 - 2 Means of Gender Related Lures in Group "fem.-masc."

| Spaced repetition | Feminine Lures | | | | | | Masculine Lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.08 | 0.12 |
| Know | 0.00 | 0.00 | 0.00 | 0.00 | 0.07 | 0.14 | 0.06 | 0.08 | 0.00 | 0.00 | 0.02 | 0.06 |
| Guess | 0.07 | 0.18 | 0.12 | 0.17 | 0.10 | 0.12 | 0.13 | 0.16 | 0.22 | 0.27 | 0.23 | 0.21 |
| No | 0.93 | 0.18 | 0.88 | 0.17 | 0.83 | 0.21 | 0.81 | 0.23 | 0.78 | 0.27 | 0.67 | 0.31 |

### Appendix C - 1 - 3 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "fem.-masc."

| Spaced repetition | Neuter Lures | | | | | | Semantic unrelated Lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | *p* | *SD* |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | | |
| Remember | 0.00 | 0.00 | 0.00 | 0.00 | 0.10 | 0.12 | 0.03 | 0.03 |
| Know | 0.00 | 0.00 | 0.14 | 0.15 | 0.02 | 0.06 | 0.05 | 0.07 |
| Guess | 0.04 | 0.12 | 0.21 | 0.22 | 0.10 | 0.12 | 0.09 | 0.13 |
| No | 0.96 | 0.12 | 0.65 | 0.19 | 0.78 | 0.26 | 0.83 | 0.18 |

### Appendix C - 1 - 4 Means of Targets in Group "masc.-neu."

| Spaced repetition | Masculine Targets | | | | | | Neuter targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.38 | 0.30 | 0.69 | 0.28 | 0.82 | 0.26 | 0.40 | 0.41 | 0.69 | 0.38 | 0.67 | 0.38 |
| Know | 0.07 | 0.09 | 0.09 | 0.12 | 0.09 | 0.18 | 0.06 | 0.12 | 0.11 | 0.17 | 0.09 | 0.12 |
| Guess | 0.07 | 0.09 | 0.09 | 0.09 | 0.05 | 0.08 | 0.09 | 0.12 | 0.09 | 0.18 | 0.13 | 0.17 |
| No | 0.48 | 028 | 0.13 | 0.17 | 0.04 | 0.12 | 0.45 | 0.40 | 0.11 | 0.20 | 0.11 | 0.23 |

**Appendix C - 1 - 5 Means of Gender Related Lures in Group "masc.-neu."**

| Spaced repetition | Masculine Lures | | | | | | Neuter Lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.00 | 0.00 | 0.02 | 0.06 | 0.08 | 0.09 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Know | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 | 0.06 | 0.00 | 0.00 | 0.05 | 0.08 | 0.02 | 0.06 |
| Guess | 0.18 | 0.16 | 0.07 | 0.12 | 0.07 | 0.12 | 0.11 | 0.17 | 0.20 | 0.14 | 0.28 | 0.18 |
| No | 0.82 | 0.16 | 0.91 | 0.18 | 0.83 | 0.18 | 0.89 | 0.17 | 0.75 | 0.19 | 0.70 | 0.20 |

**Appendix C - 1 - 6 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "masc.-neu."**

| Spaced repetition | Feminine Lures | | | | | | Semantic unrelated Lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | *p* | *SD* |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | | |
| Remember | 0.02 | 0.06 | 0.04 | 0.07 | 0.08 | 0.12 | 0.02 | 0.03 |
| Know | 0.02 | 0.06 | 0.00 | 0.00 | 0.02 | 0.06 | 0.04 | 0.05 |
| Guess | 0.18 | 0.18 | 0.13 | 0.14 | 0.17 | 0.18 | 0.10 | 0.10 |
| No | 0.78 | 0.17 | 0.83 | 0.12 | 0.73 | 0.20 | 0.84 | 0.12 |

**Appendix C - 1 - 7 Means of Targets in Group "fem.-neu."**

| Spaced repetition | Feminine Targets | | | | | | Neuter targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.27 | 0.23 | 0.70 | 0.29 | 0.90 | 0.09 | 0.44 | 0.23 | 0.70 | 0.13 | 0.70 | 0.20 |
| Know | 0.30 | 0.29 | 0.20 | 0.25 | 0.03 | 0.07 | 0.13 | 0.13 | 0.10 | 0.09 | 0.03 | 0.07 |
| Guess | 0.17 | 0.15 | 0.03 | 0.07 | 0.07 | 0.09 | 0.13 | 0.20 | 0.10 | 0.14 | 0.20 | 0.20 |
| No | 0.26 | 0.28 | 0.07 | 0.09 | 0.00 | 0.00 | 0.30 | 0.13 | 0.10 | 0.14 | 0.07 | 0.09 |

**Appendix C - 1 - 8 Means of Gender Related Lures in Group "fem.-neu."**

| Spaced repetition | Feminine Lures | | | | | | Neuter Lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Know | 0.10 | 0.21 | 0.00 | 0.00 | 0.06 | 0.09 | 0.03 | 0.07 | 0.00 | 0.00 | 0.03 | 0.07 |
| Guess | 0.10 | 0.09 | 0.20 | 0.20 | 0.17 | 0.11 | 0.00 | 0.00 | 0.07 | 0.14 | 0.27 | 0.18 |
| No | 0.80 | 0.25 | 0.80 | 0.20 | 0.77 | 0.18 | 0.97 | 0.07 | 0.93 | 0.14 | 0.70 | 0.20 |

## Appendix C - 1 - 9 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "fem.-neu."

| Spaced repetition | Masculine Lures | | | | | | Semantic unrelated Lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | *p* | *SD* |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | | |
| Remember | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| Know | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.07 | 0.01 | 0.02 |
| Guess | 0.07 | 0.09 | 0.07 | 0.09 | 0.33 | 0.11 | 0.09 | 0.07 |
| No | 0.93 | 0.09 | 0.93 | 0.09 | 0.63 | 0.17 | 0.92 | 0.08 |

## Appendix C - 2 Results of Targets

### Appendix C - 2 - 1 Tests of Within-Subjects Effects (Answer "Remember")

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Times | 8.644 | 2 | 4.322 | 1634.420 | 0.000 |
| Times *Target | 0.150 | 4 | 0.037 | 1.431 | 0.226 |
| Times*Recognition | 0.638 | 4 | 0.160 | 6.069 | 0.000 |
| Time*Targets*Recognition | 0.115 | 2 | 0.058 | 2.196 | 0.114 |
| Error (Times) | 4.416 | 168 | 0.026 | | |

### Appendix C - 2 - 2 Tests of Between-Subjects Effects (Answer "Remember")

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Intercept | 108.072 | 1 | 108.072 | 590.607 | 0.000 |
| Target | 0.116 | 2 | 0.058 | 0.318 | 0.729 |
| Recognition | 0.147 | 2 | 0.074 | 0.403 | 0.670 |
| Target*Recognition | 0.141 | 1 | 0.141 | 0.773 | 0.382 |
| Error | 15.371 | 84 | 0.183 | | |

### Appendix C - 2 - 3 Tests of Within-Subjects Effects (Answer "Know")

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Times | 0.130 | 2 | 0.065 | 3.459 | 0.034 |
| Times *Target | 0.284 | 4 | 0.071 | 3.783 | 0.006 |
| Times*Recognition | 0.152 | 4 | 0.038 | 2.023 | 0.093 |
| Time*Targets*Recognition | 0.008 | 2 | 0.004 | 0.222 | 0.802 |
| Error (Times) | 3.151 | 168 | 0.019 | | |

**Appendix C - 2 - 4 Tests of Between-Subjects Effects (Answer "Know")**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Intercept | 2.791 | 1 | 2.791 | 65.357 | 0.000 |
| Target | 0.147 | 2 | 0.074 | 1.726 | 0.184 |
| Recognition | 0.085 | 2 | 0.042 | 0.992 | 0.375 |
| Target*Recognition | 0.104 | 1 | 0.104 | 2.446 | 0.122 |
| Error | 3.587 | 84 | 0.043 | | |

**Appendix C - 2 - 5 Tests of Within-Subjects Effects (Answer "Guess")**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Times | 0.121 | 2 | 0.061 | 4.415 | 0.014 |
| Times *Target | 0.021 | 4 | 0.005 | 0.374 | 0.827 |
| Times*Recognition | 0.337 | 4 | 0.084 | 6.142 | 0.000 |
| Time*Targets*Recognition | 0.068 | 2 | 0.034 | 2.488 | 0.086 |
| Error (Times) | 2.304 | 168 | 0.014 | | |

**Appendix C - 2 - 6 Tests of Between-Subjects Effects (Answer "Guess")**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Intercept | 2.862 | 1 | 2.862 | 117.925 | 0.000 |
| Target | 0.009 | 2 | 0.047 | 1.919 | 0.153 |
| Recognition | 0.113 | 2 | 0.056 | 2.321 | 0.104 |
| Target*Recognition | 0.000 | 1 | 0.000 | 0.012 | 0.913 |
| Error | 2.039 | 84 | 0.024 | | |

**Appendix C - 2 - 7 Tests of Within-Subjects Effects (Answer "No")**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Times | 5.760 | 2 | 2.880 | 144.869 | 0.000 |
| Times *Target | 0.080 | 4 | 0.020 | 1.000 | 0.409 |
| Times*Recognition | 0.120 | 4 | 0.030 | 1.513 | 0.201 |
| Time*Targets*Recognition | 0.008 | 2 | 0.004 | 0.200 | 0.803 |
| Error (Times) | 3.340 | 168 | 0.020 | | |

**Appendix C - 2 - 8 Tests of Between-Subjects Effects (Answer "No")**

| | Sum of Squares | df | Mean Square | F | *p* |
|---|---|---|---|---|---|
| Intercept | 7.242 | 1 | 7.242 | 102.624 | 0.000 |
| Target | 0.180 | 2 | 0.090 | 1.274 | 0.285 |
| Recognition | 0.109 | 2 | 0.055 | 0.775 | 0.464 |
| Target*Recognition | 0.001 | 1 | 0.001 | 0.013 | 0.909 |
| Error | 5.928 | 84 | 0.071 | | |

# Appendix D: Results of Experiment 4

## Appendix D - 1 Means of the Main Study in Experiment 4

### Appendix D- 1 - 1 Means of Targets in Group "fem.-masc."

| Spaced repetition | Feminine Targets | | | | | | Masculine targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.38 | 0.28 | 0.83 | 0.22 | 0.71 | 0.30 | 0.34 | 0.29 | 0.65 | 0.27 | 0.74 | 0.27 |
| Know | 0.12 | 0.17 | 0.09 | 0.09 | 0.18 | 0.17 | 0.22 | 0.28 | 0.18 | 0.16 | 0.09 | 0.18 |
| Guess | 0.10 | 0.09 | 0.05 | 0.08 | 0.03 | 0.07 | 0.16 | 0.16 | 0.09 | 0.12 | 0.12 | 0.20 |
| No | 0.40 | 0.29 | 0.03 | 0.13 | 0.08 | 0.16 | 0.28 | 0.24 | 0.08 | 0.14 | 0.05 | 0.08 |

### Appendix D - 1 - 2 Means of Gender Related Lures in Group "fem.-masc."

| Spaced repetition | Feminine Lures | | | | | | Masculine Lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.04 | 0.05 | 0.08 | 0.00 | 0.00 | 0.01 | 0.04 |
| Know | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.04 | 0.03 | 0.07 | 0.00 | 0.00 |
| Guess | 0.13 | 0.16 | 0.09 | 0.09 | 0.22 | 0.24 | 0.11 | 0.12 | 0.18 | 0.30 | 0.21 | 0.20 |
| No | 0.87 | 0.16 | 0.91 | 0.09 | 0.77 | 0.23 | 0.83 | 0.18 | 0.79 | 0.30 | 0.78 | 0.22 |

### Appendix D - 1 - 3 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "fem.-masc."

| Spaced repetition | Neuter Lures | | | | | | Semantic unrelated Lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | *p* | *SD* |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | | |
| Remember | 0.00 | 0.00 | 0.05 | 0.08 | 0.04 | 0.90 | 0.00 | 0.00 |
| Know | 0.02 | 0.06 | 0.02 | 0.06 | 0.01 | 0.04 | 0.00 | 0.00 |
| Guess | 0.08 | 0.11 | 0.21 | 0.18 | 0.18 | 0.15 | 0.08 | 0.06 |
| No | 0.90 | 0.14 | 0.72 | 0.21 | 0.77 | 0.12 | 0.92 | 0.06 |

### Appendix D- 1 - 4 Means of Targets in Group "masc.-neu."

| Spaced repetition | Masculine Targets | | | | | | Neuter targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.36 | 0.26 | 0.63 | 0.26 | 0.86 | 0.20 | 0.33 | 0.281447 | 0.77 | 0.26 | 0.81 | 0.15 |
| Know | 0.19 | 0.14 | 0.21 | 0.23 | 0.13 | 0.18 | 0.18 | 0.146427 | 0.12 | 0.17 | 0.16 | 0.13 |
| Guess | 0.05 | 0.08 | 0.07 | 0.09 | 0.01 | 0.04 | 0.17 | 0.189 | 0.03 | 0.07 | 0.03 | 0.07 |
| No | 0.40 | 0.16 | 0.09 | 0.14 | 0.00 | 0.00 | 0.32 | 0.204003 | 0.08 | 0.09 | 0.00 | 0.00 |

**Appendix D - 1 - 5 Means of Gender Related Lures in Group "masc.-neu."**

| Spaced repetition | Neuter Lures | | | | | | Masculine Lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.00 | 0.00 | 0.03 | 0.07 | 0.02 | 0.06 | 0.00 | 0.00 | 0.01 | 0.04 | 0.02 | 0.06 |
| Know | 0.03 | 0.07 | 0.03 | 0.07 | 0.06 | 0.10 | 0.01 | 0.04 | 0.02 | 0.06 | 0.01 | 0.04 |
| Guess | 0.07 | 0.11 | 0.14 | 0.14 | 0.15 | 0.17 | 0.09 | 0.11 | 0.16 | 0.16 | 0.17 | 0.14 |
| No | 0.90 | 0.16 | 0.80 | 0.16 | 0.77 | 0.20 | 0.90 | 0.11 | 0.81 | 0.18 | 0.80 | 0.16 |

**Appendix D - 1 - 6 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "masc.-neu."**

| Spaced repetition | Feminine Lures | | | | | | Semantic unrelated Lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | *p* | *SD* |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | | |
| Remember | 0.00 | 0.00 | 0.02 | 0.06 | 0.00 | 0.00 | 0.00 | 0.00 |
| Know | 0.02 | 0.06 | 0.06 | 0.08 | 0.01 | 0.04 | 0.02 | 0.03 |
| Guess | 0.11 | 0.12 | 0.09 | 0.10 | 0.10 | 0.15 | 0.08 | 0.08 |
| No | 0.87 | 0.11 | 0.83 | 0.15 | 0.89 | 0.16 | 0.90 | 0.08 |

**Appendix D- 1 - 7 Means of Targets in Group "fem.-neu."**

| Spaced repetition | Feminine Targets | | | | | | Neuter targets | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.44 | 0.32 | 0.73 | 0.17 | 0.84 | 0.21 | 0.41 | 0.17 | 0.73 | 0.24 | 0.77 | 0.37 |
| Know | 0.09 | 0.12 | 0.13 | 0.11 | 0.08 | 0.12 | 0.08 | 0.185 | 0.06 | 0.10 | 0.06 | 0.12 |
| Guess | 0.17 | 0.17 | 0.08 | 0.14 | 0.05 | 0.10 | 0.16 | 0.15 | 0.07 | 0.12 | 0.13 | 0.23 |
| No | 0.30 | 0.22 | 0.06 | 0.12 | 0.03 | 0.07 | 0.35 | 0.21 | 0.14 | 0.20 | 0.04 | 0.12 |

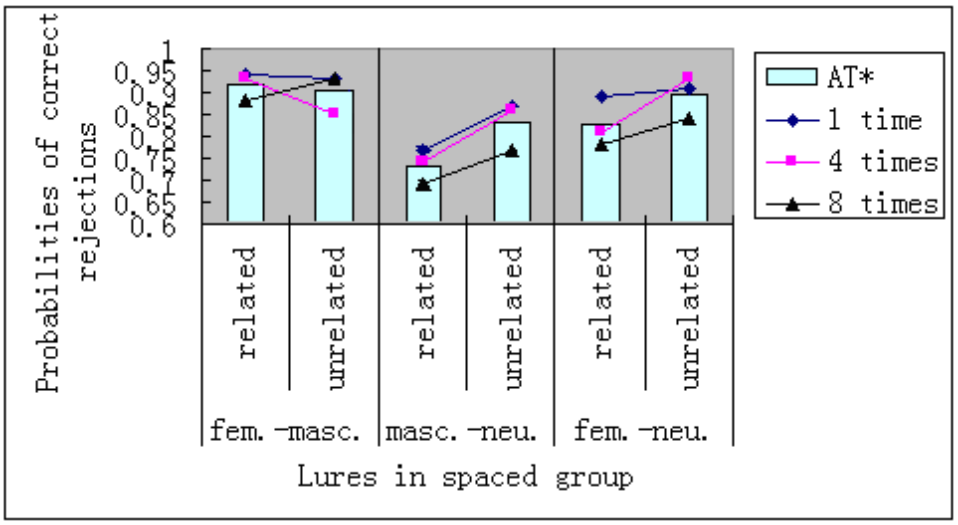
**Appendix D - 1 - 8 Means of Gender Related Lures in Group "fem.-neu."**

| Spaced repetition | Feminine Lures | | | | | | Masculine Lures | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | 1 | | 4 | | 8 | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.00 | 0.00 | 0.04 | 0.10 | 0.07 | 0.18 | 0.03 | 0.13 | 0.01 | 0.04 | 0.11 | 0.16 |
| Know | 0.00 | 0.00 | 0.04 | 0.10 | 0.10 | 0.12 | 0.00 | 0.00 | 0.02 | 0.09 | 0.06 | 0.09 |
| Guess | 0.26 | 0.20 | 0.20 | 0.14 | 0.19 | 0.19 | 0.22 | 0.22 | 0.18 | 0.19 | 0.25 | 0.19 |
| No | 0.74 | 0.20 | 0.72 | 0.18 | 0.64 | 0.31 | 0.75 | 0.29 | 0.79 | 0.20 | 0.58 | 0.22 |

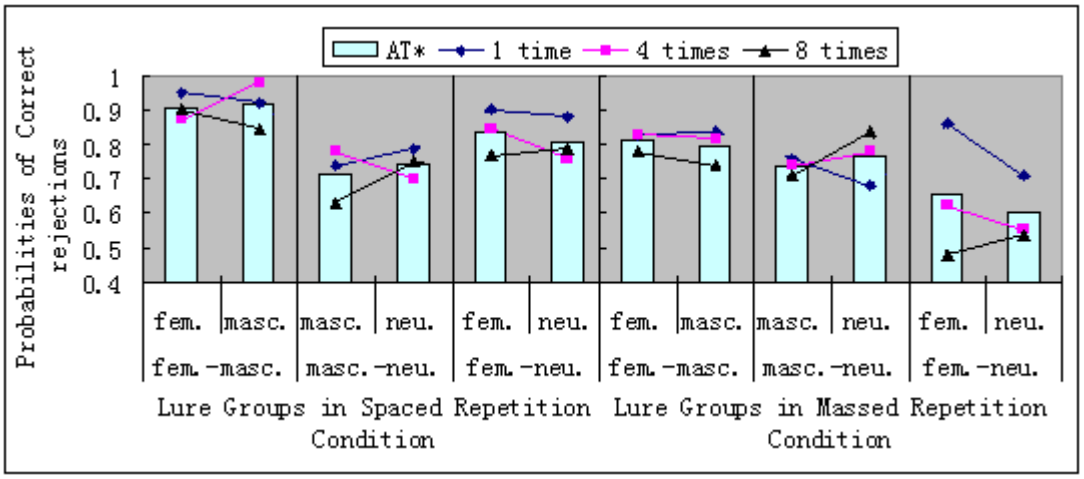
**Appendix D - 1 - 9 Means of Gender Unrelated Lures and Semantic Unrelated Lures in Group "fem.-neu."**

| Spaced repetition | Neuter Lures | | | | | | Semantic unrelated Lures | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 4 | | 8 | | | |
| | *p* | *SD* | *p* | *SD* | *p* | *SD* | *p* | *SD* |
| Remember | 0.03 | 0.13 | 0.02 | 0.09 | 0.10 | 0.15 | 0.00 | 0.02 |
| Know | 0.02 | 0.06 | 0.04 | 0.08 | 0.02 | 0.09 | 0.02 | 0.04 |
| Guess | 0.04 | 0.07 | 0.17 | 0.18 | 0.25 | 0.20 | 0.13 | 0.11 |
| No | 0.91 | 0.14 | 0.77 | 0.23 | 0.63 | 0.27 | 0.85 | 0.14 |

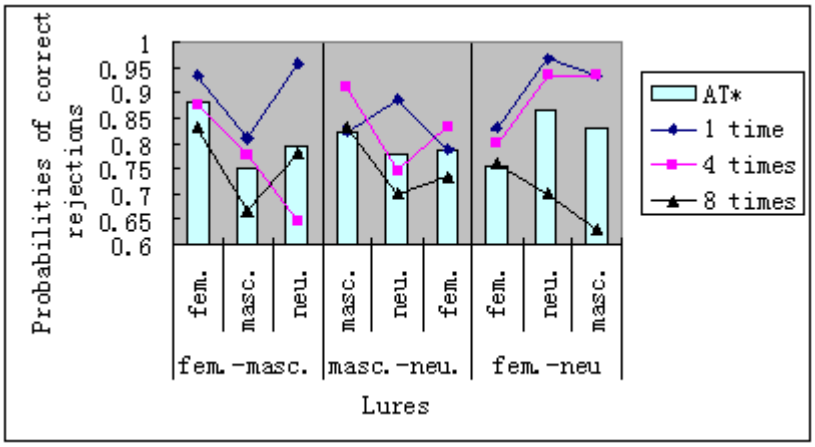
# Appendix E: Details of Figures



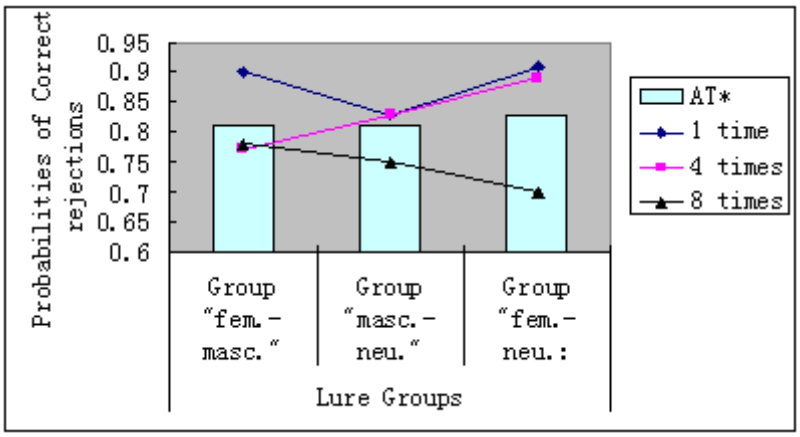

**Appendix E - 1 Comparisons between correct rejections of gender-related lures and gender-unrelated lures within different target conditions**



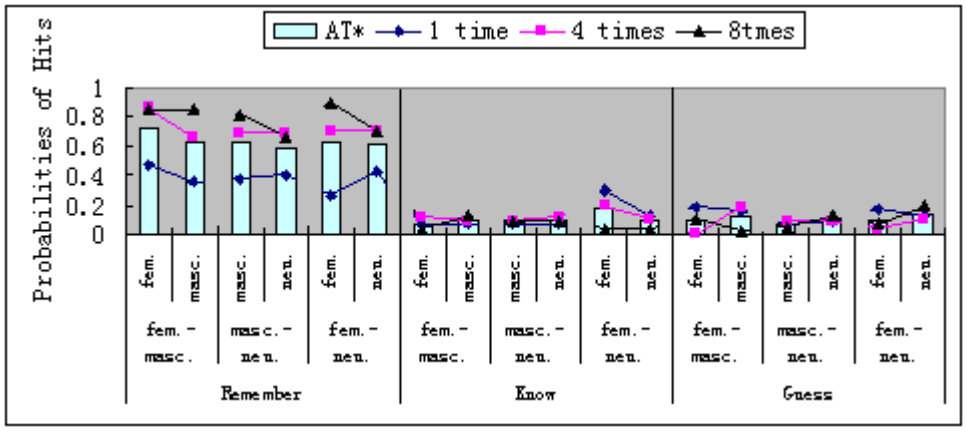

**Appendix E - 2 Comparisons between correct rejections of gender-related lures with different repetition types**



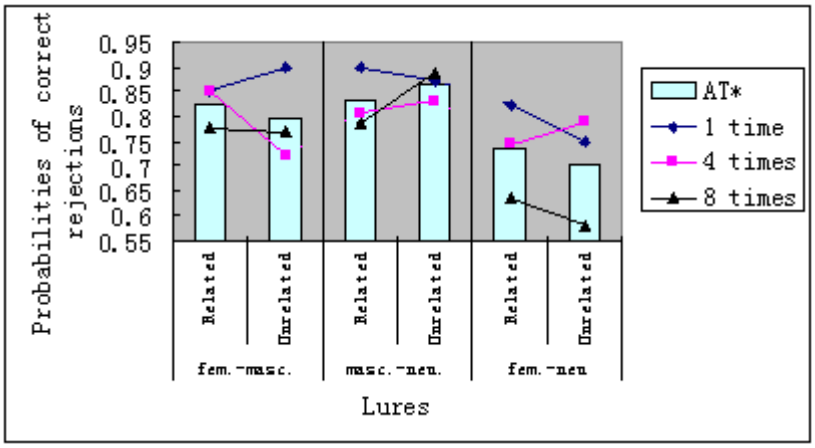

**Appendix E - 3 Comparisons of correct rejections among different lures in different target groups**



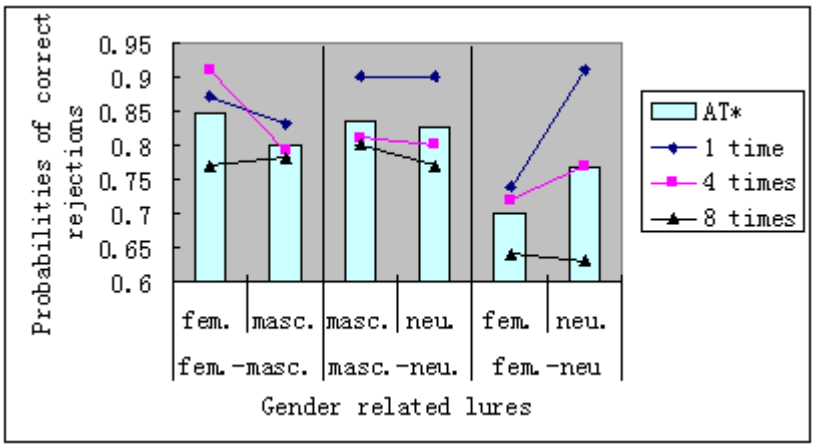

**Appendix E- 4 Comparisons of correct rejections among different target groups (Across different genders)**



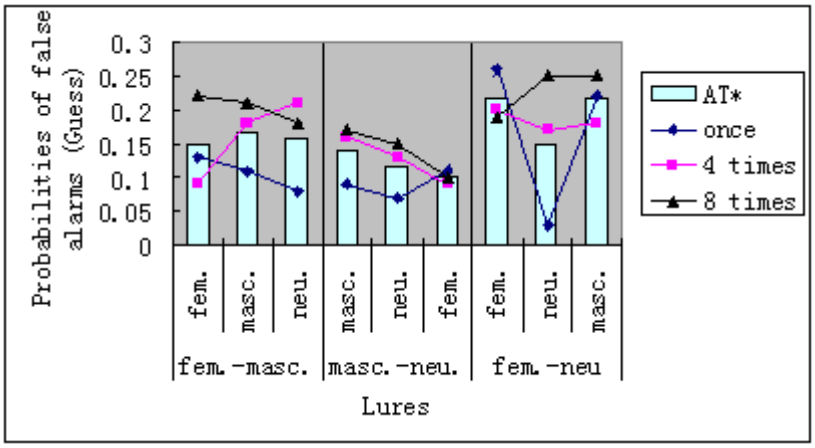

**Appendix E - 5 Comparisons of hits between targets within each target group with different response type**



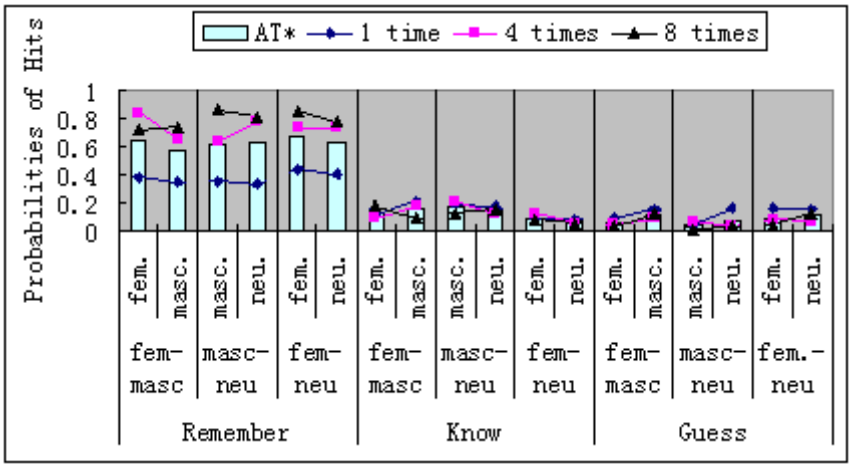

**Appendix E - 6 Comparisons between gender-related lures and gender-unrelated lures within each target group with different number of repetitions**



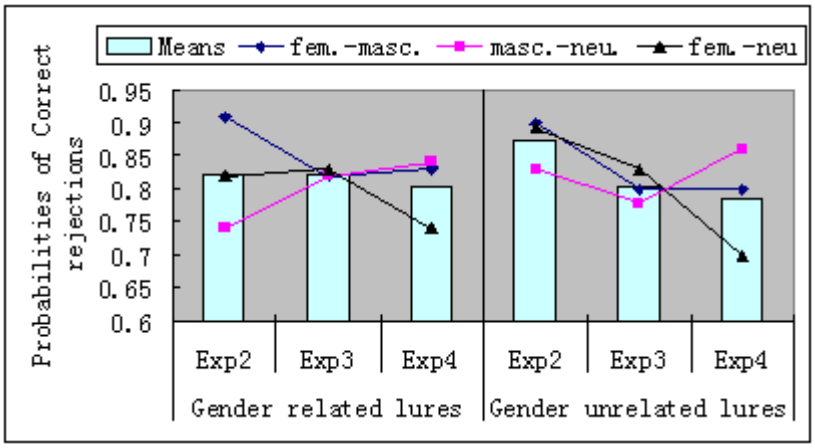

**Appendix E - 7 Comparisons between gender-related lures within each target group with different number of repetitions**



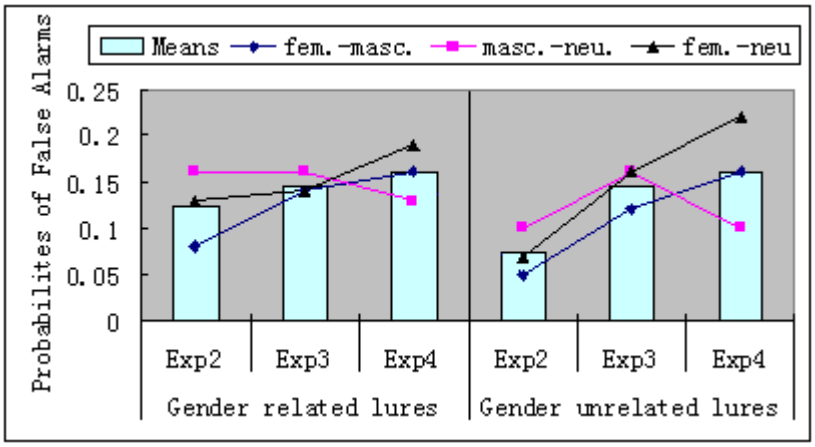

**Appendix E - 8 Comparisons of false alarms among different lures within each target group with different number of repetitions**




**Appendix E - 9 Comparisons between targets within each target group with different number of repetitions and response types**




**Appendix E - 10 Comparisons of Correct Rejections to different types of lures among Experiment 2, 3, and 4**




**Appendix E - 11 Comparisons of False Alarms (Guess) to different types of lures among Experiment 2, 3, and 4**