Master's Thesis
Advisor: Dr. Andrea Weber

Lexical Gender & Non-Native Spoken Word-Recognition

The Origin of the Gender Effect

Garance Paris
March 31, 2004

Computational Psycholinguistics
Department 4.7 General Linguistics
University of Saarland

Erklärung

Ich erkläre an Eides statt, daß ich die Diplomarbeit mit dem Titel *Lexical Gender & Non-Native Spoken Word-Recognition* selbständig und ohne fremde Hilfe verfaßt, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und alle den benutzten Quellen wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Saarbrücken, den 31. März 2004

Garance Paris

Abstract

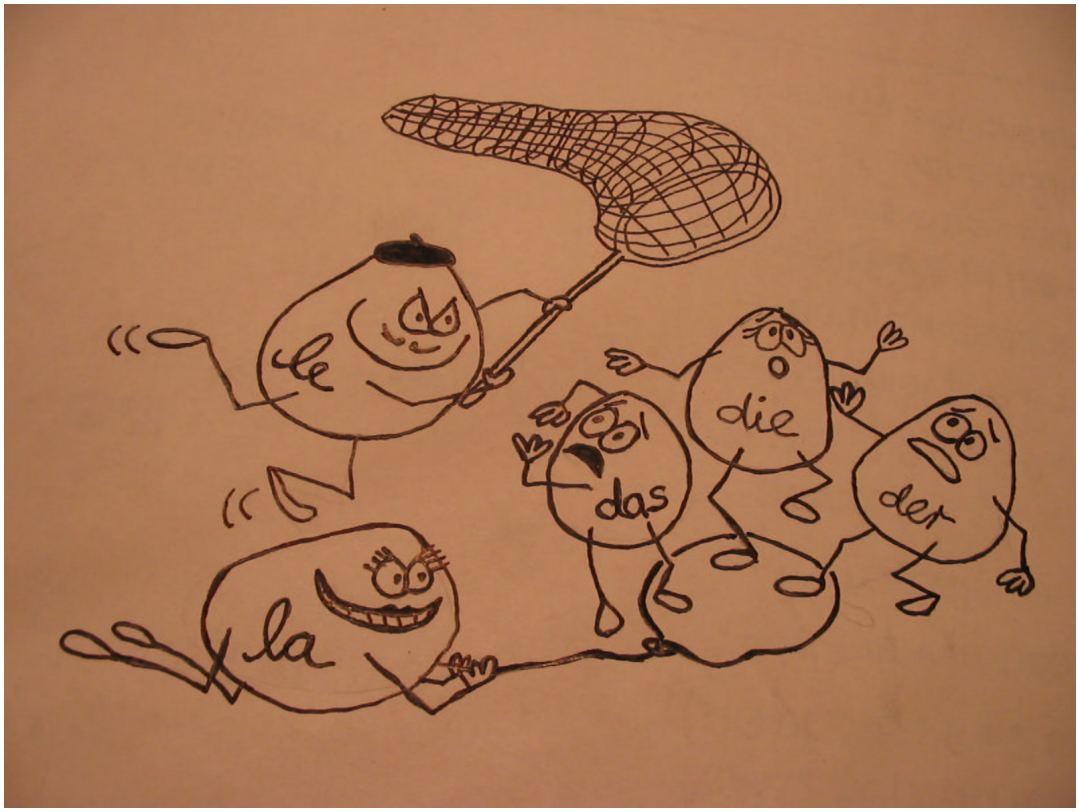
In this thesis, we investigated the impact of lexical gender on non-native spoken word-recognition. We conducted two experiments, in each of which we respectively compared proficient learners of German or French with native listeners.

Using eyetracking in visual worlds, we asked participants to click on one of four pictures. Each display contained the target, an onset-overlapping competitor and two distractors. In the instructions, the target was always preceded by an agreeing, gender-marked article. The target and competitor nouns all overlapped with their translations. After both experiments, a written questionnaire ascertained that the non-natives knew the gender of the nouns.

Experiment 1 was run in French. We compared the reactions of native Francophones and proficient Germanophone learners of French. Half the items were based on same-gender target-competitor pairs, whereas the other half were based on pairs which differed in gender in French, but not in German. In the same-gender pairs, native listeners showed competition (longer fixation of competitor than distractors), but in the different-gender pairs, they did not: when appropriate, they made use of the article's gender to constrain the set of lexical candidates, as in Dahan et al. (2000). Non-natives listeners, however, showed competition in both conditions, thereby revealing that they could not use gender as the natives did.

In Experiment 2, we used the same materials in German to test Germanophone natives and Francophone learners of German. Here, on the surface, all noun pairs had the same gender; the only difference between same-gender and different-gender pairs lay in the underlying French translations of the competitors. As expected, Germanophone natives showed competition in both conditions, but interestingly, the Francophones did not show competition in the different-gender pairs, suggesting that they used French gender while listening to German.

Such a use by non-native listeners of their native gender categories in foreign language processing thus allows us to affirm that, in native processing, the gender effect (i. e. faster recognition of a spoken noun after congruent gender-marking) is due to the use of grammatical gender categories, and not to the surface forms of articles and nouns, which vary very much from language to language.



Native gender categories capture gender in foreign language
(see Section 4.2)

(Illustration by Ruth Kusterer)

✠
JMJ

Acknowledgments

. . . reddite quae sunt Caesaris Caesari et quae sunt Dei Deo.[1]

The months I spent writing my Master's thesis have been an enjoyable conclusion of my university studies in Saarbrücken. On the eve of handing in this thesis, I am glad to have many people to thank for their help and support throughout these years, and throughout this project in particular. First of all, there is one whom I would like to thank for my landing here in the first place: Lord, thank you for letting me study computational linguistics in such a stimulating place as Saarbrücken!

Then I would like to thank Professor Matthew Crocker for all the support and encouragement I received from him during my studies in Saarbrücken, for having created such a pleasant working atmosphere in his department, and for providing the financial support for running subjects.

Most of all, however, I am deeply grateful to my advisor, Andrea Weber, for her patience with me over these months, and for her helpfulness from beginning to end. Having studied computational linguistics, not psycholinguistics, I lacked many basics when I started out on this project, and she had to teach me theory, eyetracking, statistics, writing up experimental results, and much more. I could always knock on her door anytime to ask questions, and her support was fundamental to my completing this project.

Many thanks also go to the other people in the department who helped me at various points: Thorsten Kaske, who programmed the "mouseclick" and analysis softwares, Alissa Melinger, who answered my questions on English and Italian, and our two secretaries, Claudia Verburg and Magdalena Mitova, who smoothed over many details, especially the paying of the participants. Thanks also to Gudrun Schuchmann and Caren Brinckmann (from the phonetics department), who lent us their voices for the German experiment, and to Steffi Schwabe and Katja Pfahler, the student assistants who ran some of the German subjects in the German experiment.

[1] ". . . render to Caesar the things that are Caesar's; and to God, the things that are God's" (Ma, 22:21). English translation from the Douay-Rheims Bible.

I am also very grateful to the participants in the experiments, especially to the many who are not students of the University of Saarland, but came from further away just to participate—and, what's more, often did not even want money in compensation (although they did take Belgian chocolate)! Moreover, I am also thankful to those who helped me by advertising for the project: Peter Feid, Mike Vaughan, and others.

Of course, my friends at college were also a big support throughout these years, and especially during the past months. In particular, I would like to thank "Swiss Ruth" (aka Ruth Kusterer) and "Punk Ruth" (Ruth Fuchss), as well as Ulrike and Sebastian "Padowein", with whom I had many a good laugh during lunch time, and who have been such wonderful fellow students, from which I received much help and learned a lot over the past few years. I am also very much indebted to Christian Dressler, who taught me to use Linux and made my computer work. Very special thanks go to my two flatmates during these years: Stephanie Köser, who put up with me during four years and taught me German, cooking and more, and Maria-Antonietta Baldi, a real Italian sunshine. If it had not been for you all I would probably have given up on computers and on the German language a long time ago!

Finally, there are two other persons which I would also like to thank—without whom I would not even be here, let alone be able to write a Master's thesis: they are my two very loving parents: Thanks Mom and Dad! I love you :-))

Contents

List of Figures and Tables

1 Introduction

> *Every noun has a gender, and there is no sense or system in the distribution; so the gender of each must be learned separately and by heart [...] In German, a young lady has no sex, while a turnip has.*
> (*The Awful German Language*, by Mark Twain)

Although absent from many languages, lexical gender is omnipresent in many others, "emerging in virtually every sentence one might want to construct" (van Berkum, 1996, p. 1). Often, different forms of the article are used depending on the gender of the following noun. Children learn what article belongs before which noun with astonishing ease in their mother-tongue, and adults hardly make any mistakes in their choice of articles, but for non-native speakers, gender is very difficult to learn—as illustrated by the Mark Twain citation above.

Moreover, at first sight, it is not even obvious what communicative function gender may serve, making it all the more annoying to learn in a foreign language because it seems relatively useless. However, native listeners have no trouble picking up gender mistakes and correcting them very casually, thus proving that they do somehow notice the gender of an article. This observation may be the reason why psycholinguists started investigating possible functions of gender, eventually showing that, among other benefits, one possible advantage of gender may be that it can be used by native listeners to facilitate spoken word-recognition, because a gender-marked article restricts the search space within the mental lexicon to congruent nouns (see Sections 1.3 and 1.5).

Intrigued by non-natives' frequent difficulties with the production of gender-marked utterances, Guillelmon & Grosjean (2001) then researched the sensitiveness of second-language learners to gender in their foreign tongue, and found that, in spoken word-recognition, English learners of French were insensitive to French gender, and thus unable to profit from it. In reading, Sabourin (2003) also showed that non-native listeners do not process gender similarly to natives: she demonstrated that, when asked to tell whether a sentence was grammatical or not, non-natives' neurophysiological responses were different from those

of natives, and that the accuracy of their answers was lower. Furthermore, she also found a strong influence of the participants' mother-tongue: when the mother-tongue had a gender system similar to that of the experimental language, the participants did better than participants whose mother-tongue had lexical gender but where the categories were very different, and the latter did again better than participants whose mother-tongue did not have gender at all. This Sabourin explained by saying that non-native listeners attempt to transfer their mother-tongue categories to their foreign language.

In this thesis, we also asked ourselves what the impact of native gender categories on foreign language processing would be, but applied this question instead to the domain of spoken word-recognition. Indeed, it is known that at other stages of processing, such as on the phonetic level, non-natives are not sensitive to input in the same way as native listeners, with consequences on the way they recognize speech in their foreign language (see, for example, the results of Weber & Cutler, 2004, described on page 18). Given these observations on the phonetic level and the results of Sabourin (2003), we thus wondered whether non-native listeners would be able to learn to use the gender categories of their foreign language, or whether they might transfer the gender categories from their mother-tongue and apply them also in non-native spoken word-recognition.

Additionally, in psycholinguistics, the observation of non-native processing is especially interesting when it can shed light onto more general mechanisms, including those at work in native processing. As we will see, our results also have implications for native spoken word-recognition, because the interference of mother-tongue and foreign language which will be our conclusion in Section 4.3 also offers an explanation of how gender is used by native listeners.

Plan of the Thesis

This thesis is organized as follows. In the remainder of this introduction we will first briefly introduce the basic ideas behind current models of spoken word-recognition, and explain why it is generally thought that non-native listeners have to deal with a much larger search space through the lexicon in spoken word-recognition (Section 1.1). Following this, we will then give a linguistic review of the concept of lexical gender and its potential functions in human language processing (Section 1.2), as well as an overview of a number of offline

experimental studies of the 'gender effect' (i. e. the facilitatory effect on noun recognition of a preceding congruent gender-marked word, and the inhibitory effect of an incongruent gender-marked word) on spoken word-recognition (Section 1.3). In particular, we will show that these studies do not allow us to tell at what point of spoken word-recognition gender has its impact. We will then introduce the methodology of eyetracking in visual worlds (Section 1.4), which allowed Dahan et al. (2000) to determine the precise point in time at which gender acts on spoken word-recognition (Section 1.5). Finally, we will expose the question of the origin of the gender effect, which was broached by Dahan et al., and explain how we proposed to approach it by studying the behavior of non-native listeners (Section 1.6).

The following two chapters will then present our experiments. We conducted two versions of the same experiment, the first with French instructions (Experiment 1, Chapter 2) and the second with German instructions (Experiment 2, Chapter 3). In both cases, the experiment was run with both native (Experiments 1a and 2a) and non-native listeners (Experiments 1b and 2b), and the results were compared.

In the General Discussion, we will first give an overview of our experimental results, and then see how our Francophone natives in Experiment 1a replicated the study by Dahan et al. (2000; Section 4.1). Following this, we will discuss the data we obtained with non-native listeners, and show how native gender categories appear to take over the gender categories of the foreign language (Section 4.2). We will then come back to the origin of the gender effect, of which we will show that it must be grammar-based, also in native processing (Section 4.3). Finally, we will conclude with a summary of our theoretic conclusions and some incentive for future work (Section 4.4).

1.1 Native and Non-Native Spoken Word-Recognition

All currently accepted models of spoken word-recognition assume that during word-recognition, a set (or 'cohort') of lexical representations consistent with the phonetic input are simultaneously activated in the mental lexicon. Members of this set then compete for recognition, each mismatch between the phonetic

input and a potential candidate reducing its activation. How long it takes to recognize a word depends on the competition set: the more candidates and the tighter the competition, the slower the process (Norris, McQueen, & Cutler, 1995; Vroomen & de Gelder, 1995). Recognition is achieved when only one word remains, or one word is much more highly activated than all others (for reviews of word-recognition theories and models, see Tyler & Frauenfelder, 1987, and Cutler, 1995).

Thus, depending on the exact assumptions made by a given model, a native Francophone listener hearing the beginning of a noun starting with /k/ might activate all 2326 French nouns which, according to New, Pallier, Ferrand, & Matos (2003), begin with that sound, among others *camion* (/kamjõ/, English: "truck"), *képi* (/kepi/, "policeman's cap"), *cassette* (/kasɛt/, "tape"), and *queue* (/kø/, "tail"). A following /a/ would then reduce the competition set to only 623 nouns, a following /s/ to 34, and this would go on until only one noun remained available for selection: in this case, as soon as a following /ɛ/ was heard, the listener would know that the word was *cassette*, even before the end of the noun, although in many cases, a word's uniqueness point may not be reached before its end, or before the beginning of the following word.

Between the various models of spoken word-recognition, there are, however, differences in the precise way in which the competition set is defined: in early versions of the Cohort Model (Marslen-Wilson & Welsh, 1978), only the onset of a word activated candidates for recognition, whereas in later versions (Marslen-Wilson, 1987) and in other models (TRACE: McClelland & Elman, 1986; SHORTLIST: Norris, 1994), any part of the input may activate candidates. This is supported by evidence from various experimental studies (see the overview by Cutler, 1995, on pp. 102–103), which have shown that rhyming words may also be activated by the phonetic input, although more weakly: *beaker*, for example, was shown to activate not only *beetle* but also *speaker* (Allopenna, Magnuson, & Tanenhaus, 1998) (see page 1.4).

Non-native speech processing is generally assumed to proceed similarly, but with the interesting peculiarity that listeners seem unable to deactivate the vocabulary of their native language, even in a monolingual situation in which they are required to communicate entirely in a non-native language. Using eyetracking in visual worlds, this was demonstrated by Spivey & Marian (1999, see also Marian & Spivey, 1999), who showed that Russian citizens living in the

United States (i. e. expatriates using a foreign language on a daily basis) tended to activate the Russian word *марка* (/marka/; English: "stamp") when hearing the beginning of the English word *marker*, and vice-versa. Following this, Weber & Cutler (2004) also showed that, in an English experiment, Dutch university students living in the Netherlands, which were highly proficient in English, activated Dutch competitors such as *kist* (English: "chest") when hearing the first sounds of *kitten*.

Unfortunately, this inability to deactivate the native lexicon means that the competition set is much larger in non-native than in native spoken word-recognition, thus leading to larger competition effects and slower recognition. What's more, other factors also seem to contribute to increase the size of the competition set: the results of Weber & Cutler (2004) also suggested that native phonemic categories have a tendency to "capture" non-native phonetic input, thereby potentially causing inappropriate competitor activation. For example, Dutch native speakers listening to English were proven to map the English /æ/ onto the Dutch phoneme /ɛ/ for the purpose of lexical access (probably because /ɛ/ is the closest phoneme in their mother-tongue resembling the English /æ/), and thus activated *pencil* as a competitor when hearing the beginning sounds of *panda* (by comparison, native listeners of English activated only *panda*, but not *pencil*).

This results in a fairly gloomy view of non-native spoken word-recognition: people listening to a foreign language seem to be doomed to making more effort in order to recognize words in the fast-flowing speech surrounding them. Moreover, pondering these results raises another question: what about other levels of processing? Is this "capturing" of foreign language input by native categories limited to the phonetic level, or does it extend to, say, morphosyntactic categories?

Just as some models of language processing in general assume modularity of the different levels while others posit interactivity, some models of spoken word-recognition assume that recognition is determined only by the phonetic input, whereas others make way for an influence from syntax and semantics (see Frauenfelder & Tyler, 1987). In this latter context, one might thus ask what the influence of morphosyntactic features coded in the lexicon would be—and again, here there might be differences between native and non-native listeners.

In this thesis, we explored the influence of one such feature, namely lexical gender, comparing the behavior of native listeners with that of non-native listeners. As Sections 1.3 and 1.5 will show, it is nowadays known that gender has an influence on spoken word-recognition in native listeners, which, in some cases, can make use of gender information to reduce competition. However, it remains as yet unclear whether non-natives are able to use gender in the same way, or more generally, what the influence of gender categories in the mother-tongue might be on non-native spoken word-recognition.

Indeed, although lexical gender is absent from many languages, those languages which *do* carry it frequently have at least some gender categories in common, because all gender systems seem to have originated from ancient male/female or animate/inanimate classifications (Corbett, 1991, pp. 20, 311–312; van Berkum, 1996, p. 19). However, at the same time, the coverage of each category is bound to be different in each language, because the gender systems of the various languages have evolved in very different ways.

Together, these two aspects of gender systems might allow us to observe possible morphosyntactic "capturing" effects in non-native spoken word-recognition: are non-native listeners able to use the gender information of the foreign language they are listening to, or does gender in their mother-tongue overrule the gender categories of the non-native language? If this were the case, this interference of native categories might lead to an incorrect initial set of lexical candidates, either by ruling out words which should not be ruled out or by allowing the unnecessary activation of words which should not be activated, because they would complete the utterance in an ungrammatical way. The ultimate consequence of this would then be "delayed recognition at best and misperception at worst" (Grosjean, Dommergues, Cornu, Guillelmon, & Besson, 1994).

1.2 Lexical Gender

"Are you trying to tell me that a table has a sex in French?!?"
(Remark by an English-speaking student of
French to her teacher during the first lesson)

Although an unknown phenomenon in many languages around the world, the grouping of common nouns into classes according to the way they influence "the behavior of associated words" (Hockett, 1958, p. 231) is pervasive in other languages. The resulting categories, which may or may not correspond to a

real-life distinction of sex, are ordinarily referred to as *genders*, by derivation from the Latin word *genus* for "class" or "sort" (Corbett, 1991, p. 1).

In order to make out what genders a given language has, the determining criterion is agreement: not all possible classifications of the nouns of a language are called genders, but only those which distinguish between classes of nouns requiring a particular form of the words associated with them.

Indeed, only some gender-marking languages allow part of their nouns to be assigned to one gender class by taking a look at their form. These are said to have *overt* gender, as opposed to languages with *covert* gender, in which the gender of most nouns cannot so easily be guessed (Corbett, 1991, p. 62). In all cases it is always agreement of the words "associated" with the nouns (which may include adjectives, many types of determiners and pronouns, participles and verbs—even adverbs, adpositions and conjunctions; Corbett, 1991, pp. 4 & 106–114; van Berkum, 1996, p. 24), and not the form of the nouns themselves, which determine what gender classes a language has and to which gender a given noun belongs.

For example, in French, which has two genders, masculine and feminine, the article and possible adjectives related to a noun will often betray its gender, which otherwise is not always easily recognizable from the form of the noun itself. Thus, for instance, in the following two noun phrases, nothing in either the phonological or the written forms of the French nouns *bateau* and *maison* indicate to the unknowing (child or non-native) reader or listener that they are respectively masculine and feminine, but the preceding articles, *le* for a masculine noun and *la* for a feminine noun, as well as the accompanying adjectives, here respectively *grand* and *grande*, foretell the gender of the noun about to follow:[2]

*l***e**	*grand*$_{\emptyset}$	*bateau*	*l***a**	*grand***e**	*maison*
/lə	grɑ̃	bato/	/la	grɑ̃d	mɛzõ/
Art-Masc	Adj-Masc	$N_{[masc]}$	Art-Fem	Adj-Fem	$N_{[fem]}$
'the big boat'			'the big house'		

[2]In this thesis, I have used small capitals (subscript or not) to indicate gender-marking on words related to nouns, whereas a subscript gender indication between brackets indicates the covert gender of the noun itself.

Most linguists agree that gender is a fixed property of each individual common noun, which must be coded in the lexicon together with the noun, but this view is not shared by all: after reviewing how various languages around the world assign gender to their nouns, Corbett (1991), for one, concludes that it would actually be possible for listeners to determine the gender of a noun on the fly according to some set of rules based, depending on the language, on semantics, morphology or phonology. This would then render individual storage of the gender information for each noun unnecessary, and would explain the high degree of consistency with which borrowings from other languages or invented words are classified by native speakers.

This latter opinion, however, does not seem to be supported by experimental evidence. Indeed, gender assignment systems, whether they rely on semantics or on the form of noun, must in all cases necessarily make use of the complete word. For example, in those proposals cited by Corbett (1991) of assignment rules to predict the gender of French nouns depending on their phonological form (achieving approximately 85 % correct; Bidot, 1925; Meľčuk, 1958; Tucker, Lambert, & Rigault, 1977), noun endings build the basis for most assignment rules. As we will see in Section 1.5, there is experimental evidence by Dahan, Swingley, Tanenhaus, & Magnuson (2000) to suggest that listeners are aware of the gender of a noun as soon as the noun begins to be activated in the mental lexicon by its first few sounds.

Consequently, we may safely admit the idea that, although languages may have gender assignment systems which are used by listeners to classify new words, there must be a difference between one-time gender assignment and gender knowledge of previously encountered words, which is retrieved from the lexicon to be used in everyday speech processing.[3]

At first sight, gender seems to uselessly complicate a language, since many languages do very well without it. Especially learners of a foreign language which has gender find gender very hard to cope with, and often, even after years of learning, non-natives continue to make a substantial number of mistakes in the production of gender-related word-forms. Moreover, as far as comprehen-

[3]This is the reason why I have preferred to use the term 'lexical' gender instead of its alternative 'grammatical gender' which is used by many authors, who thereby stress the agreement aspect instead of the lexical one.

sion is concerned, van Berkum (1996, p. 9) mentions that the mere existence of gender in a language "does not mean that those at the receiving end benefit from it in any way"—gender could very well be the result of a historical "accident" in a language, without being of any use for processing.

However, giving the issue some more thought does reveal a number of ways in which gender might turn out to be useful in language processing (for an overview, see van Berkum, 1996, p. 18–22). The most obvious seems to be that gender can help disambiguate between possible referents for anaphors, since many languages mark gender on pronouns. For example, in the following story, depending on the gender of the French pronoun in the second sentence (*le* for a masculine referent and *la* for a feminine one), the mischievous child will later find either only melted chocolate in his pocket, or a praline in a box which he can eat:

Le chocolat$_{\text{i}}$ *était dans une boîte*$_{\text{j}}$.
N$_{[masc]}$ N$_{[fem]}$
'The chocolate was in a box.'
Jean le$_{\text{i}}$*/la*$_{\text{j}}$ *cacha dans sa poche.*
Pron V
'John hid it in his pocket.'

Another case in which gender can be helpful is the disambiguation of some types of syntactic constructions. Take, for example, sentences with an NP-(Prep)-NP-RC complex, such as the one below. In English, the relative clause can be attached to either of the noun phrases, and is preferentially attached to the second one (Kempen, 1996; Cuetos, Mitchell, & Corley, 1996), but in German, where the relative pronoun is marked for gender, only one possibility remains: if the relative pronoun is neuter, then the relative clause must be attached to the first noun, whereas if it is masculine, it will automatically be attached to the second noun.

Hans kaufte das Buch$_{\text{i}}$ *des Schrifstellers*$_{\text{j}}$*, das*$_{\text{i}}$*/den*$_{\text{j}}$ *er mag.*
N$_{1[neut]}$ N$_{2[masc]}$ Pron
'John bought the book of the author which he likes.'

During the process of comprehension, there is however a more fundamental way in which gender may assist the reader/listener, long before the disambigua-

tion of ambiguous syntactic constructions or the resolution of anaphors takes place: at the lexical level, gender may indeed facilitate word-recognition, since the gender-marking words that are "associated" with a noun, when they precede it, partly predict what nouns are allowed to follow, by hinting at the gender of the upcoming noun.

If this were the case, it might partially explain the apparent arbitrariness with which objects belonging to the same semantic field are sometimes assigned to a gender class in some languages: Zubin & Köpcke (1981, 1986), for example, mention that in German, which has three genders, the words for "mouth", "nose" and "eye" each belong to a different gender—*der Mund* (masculine), *die Nase* (feminine), *das Auge* (neuter)—and report that among kitchen utensils, they found 41 % of masculine nouns, 42 % of feminine nouns and 17 % of neuter. By dividing nouns into gender classes in such an erratic fashion, languages might by striking a balance between motivated gender assignment, which would reduce the amount of memory required, and, at the other extreme, an arbitrary uniform distribution of gender in a given perceptual field, which, although leading to high uncertainty for the learner, would maximize the information gained from gender cues preceding a noun, thus increasing one of the possible communicative functions of gender.

1.3 Offline Studies of Gender & Word-Recognition

In reading, a number of studies have shown that after a congruent gender-marked word, native listeners recognize the following noun more rapidly (Gurjanov, Lukatela, Lukatela, Savic, & Turvey, 1985; Schmidt, 1986; Colé & Segui, 1994), although other studies proved inconclusive (van Berkum, 1996).

However, it must be remembered that the results from reading studies do not necessarily generalize well to spoken word-recognition, because we do not read linearly (from left to right in most languages), but instead perceive whole words or chunks of words with one gaze and recognize them from their overall shape, so that readers might be able to use gender information available from the ending of a word, as has been shown by Colé, Pynte, & Andriamamonjy (2003; see the remark on French nouns and their gender-transparency on p. 21).

In speech, the first to investigate gender effects were Grosjean, Dommergues, Cornu, Guillelmon, & Besson (1994). Using gating and lexical decision, they found a strong facilitatory effect when a French adjective-noun nominal phrase (containing a phonetically invariable adjective) was preceded by an agreeing article as opposed to when the article was absent. In particular, the candidates for completion which were proposed by participants at different gates were interesting: when the nominal phrase was not preceded by an article, after hearing only the adjective, the participants proposed words belonging to both French gender classes, whereas when the nominal phrase was preceded by an article, only nouns agreeing with the article were mentioned. Moreover, as would be expected since gender divides the French lexicon in two more or less equal halves (54 % masculine, 41 % feminine and 5 % ambiguous nouns; percentages computed on the basis of New et al., 2003), the stimulus word was proposed earlier when gender information was available and there were fewer candidates proposed overall.

In Italian, Bates, Devescovi, Hernandez, & Pizzamiglio (1996) subsequently showed that gender priming effects could be observed with various types of tasks (word-recognition, gender monitoring and grammaticality judgment), independently of whether the participant's attention was consciously directed to gender, and that they involved both facilitation and inhibition: reactions to nouns agreeing with the preceding article were faster than to neutral control nouns, while reactions to gender-mismatching nouns were slower. Following this, the lexical decision experiments by Jakubowicz & Faussart (1998) went on to reveal that, in French, the gender effect was weaker when a determiner alone (article, demonstrative or possessive) preceded the noun than when a determiner and a phonetically marked or unmarked adjective preceded the noun. Finally, using cross-modal priming, Spinelli & Alario (2002) showed that, in French, after a gender-marked article, only homophones matching the gender of the preceding article were activated.

Moreover, apart from these studies on how native listeners process gender, there is also one study on how non-natives deal with the phenomenon: Guillelmon & Grosjean (2001) used auditory naming to examine how "early" English-French bilinguals (those who acquired two languages before adolescence; we will refer to them simply as 'bilinguals') react to gender-marking when processing French, and compared this with the reactions of non-native listeners

(i. e. English natives who had learned French during adolescence or later and still became quite proficient in this language). They observed that the bilinguals showed clear facilitation and inhibition effects, but that the non-natives appeared to be totally oblivious to gender-marking, be it congruent or incongruent.

Coming back now to the data from native listeners, the various facilitatory and inhibitory effects observed in all gender studies appear to suggest that the gender effect takes place at the pre-lexical stage, before word-recognition is achieved. This, of course, presupposes some kind of interactivity in the process of spoken word-recognition, since morphosyntactic information such as gender, at the very least, must be allowed an influence. Although Bates et al. (1996) note that inhibitory effects in grammatical priming once were mainly interpreted as being controlled, strategic and/or post-lexical, in the case of the gender effect on spoken word-recognition, it is also possible that inhibitory effects would be due to the early activation of a wrong competition set which originally did not contain the correct target.

Thus, as mentioned by Grosjean et al. (1994), gender marking seems to be another factor (next to frequency, length, uniqueness point, neighborhood size and frequency, syntactic, semantic and pragmatic constraints, etc.) which accounts for the time it takes to identify a word, and it does not seem to be the case that words are first identified irrespective of their gender, with a post-lexical syntactic check then ensuring that agreement is respected (however, see Friederici & Jacobsen, 1999, for arguments opposed to this view).

This does not mean, of course, that a later syntactic check is completely excluded: in more complex constructions in which more than a few words intervene between a gender-marked element and the noun it modifies, it is possible that some kind of late syntactic check still takes place, although, due the distance, the gender-marked element may not necessarily have an influence on noun-recognition. This might be the case, for example, in French, when a detached attributive adjective precedes its noun by many words such as in *"Construit***e**$_{\text{FEM}}$ *au début du siècle, cette table*$_{[fem]}$*..."* (English: "Built at the turn of the century, this table..."; example taken from Grosjean et al., 1994). However, it must be pointed out that such sentences are more common in the written medium, occurring only rarely, if anything, in spoken language, and

that there is no evidence that the above experimental results also generalize to such turns of phrase.

As pointed out by both Grosjean et al. (1994) and Bates et al. (1996), the simplest explanation for a lexical effect such as the one which emerges from the above evidence would seem to be spreading-activation within the lexicon, possibly within connectionist models of spoken word-recognition such as TRACE (McClelland & Elman, 1986) or SHORTLIST (Norris, 1994): when a gender-marked word is recognized, it would then activate all nouns in the mental lexicon which share the same gender.

This however, would mean that gender-based activation would be spreading in a structure-independent manner, independently from syntactic relations, and Bates et al. (1996) argue that this is unlikely because, in gender-marking languages, sometimes other nouns belonging to other gender classes may intervene between a gender-marked word and the noun it depends from. For example, in the following German noun phrase, the neuter article announcing the word *Buch* (English: "book") is separated from it by another modifying noun phrase containing a feminine noun:

das_{i}	der_{j}	$Lehrerin_{\mathrm{j}}$	*gehörende*	$Buch_{\mathrm{i}}$
Art-NEUT	Art-FEM	$\mathrm{N}_{[fem]}$		$\mathrm{N}_{[neut]}$
the	the	teacher	belonging	book

'the book belonging to the teacher'

Again however, it must be noted that such structures are more complex than the materials used in the above-mentioned studies, and that there is nothing to allow us to simply assume that gender-priming also occurs in such complex structures. Moreover, it has been shown that there are connectionist architectures able to deal with embedding (see, for example, Elman, 1991), so that a spreading-activation mechanism within a connectionist framework would not necessarily have to be ruled out at this point.

If we admit that the gender effect takes place at the pre-lexical level, the next question then is: at what point during spoken word-recognition? Indeed, Tyler & Frauenfelder (1987), who attempt to decompose the process of spoken word-recognition into several steps, distinguish between an 'activation' phase and a 'selection' phase, the former referring to the stage at which the original set of competing lexical representations are activated and the latter to the process

during which this competition set is gradually reduced until only one entry remains.

Does gender, then, hinder some nouns from being taken up into the competition set in the first place, or are nouns which do not match a preceding gender cue activated alongside others, but then rapidly eliminated during competition? In order to answer this question, we need to take a look at results obtained with another experimental paradigm, which offers a higher time-resolution and allows us to observe exactly what competing lexical representations people consider over the course of time.

1.4 Eyetracking in Visual Worlds

Such an insight into the spoken word-recognition process is offered by head-mounted eyetracking in visual worlds. In this paradigm, small cameras observe and record the participants' eye-movements while they are instructed to either manipulate real objects lying on a table before them (as, for example, in Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), or to mouse-click on pictures of objects on a computer display (and, in some cases, to move around the objects on the screen; see among others Allopenna et al., 1998). Such a setup is referred to as a 'visual world'.

The intuition is that, since the participants are required to do something with the objects set out before them, they will first need to identify what objects they are instructed to manipulate. This, of course, involves understanding the instructions and recognizing the words therein contained. Thus, according to models of spoken word-recognition, the participants are at first expected to activate all words which share the same beginning as the noun in the spoken instruction—possibly taking other influence factors into account—and, as the speech input unfolds, to gradually restrict their attention to those among the displayed objects which best correspond to the instructions.

This is supported by the work of Tanenhaus et al. (1995), who demonstrated that spoken word-recognition is an incremental process, and that people look at relevant objects shortly after hearing only partial word information: when people were asked to click on an object (the 'target'; say, a *candle*) in the presence of another object with a similar onset (termed the 'competitor'; for example a *candy*), it took them longer to identify the target than when it was surrounded only by objects whose onset did not overlap (which are called 'distractors'),

presumably because in the first case, they had to wait until more acoustic information came in that distinguished the target from the competitor, whereas in the second case, the beginning of the noun was sufficient. More precisely, saccadic eye-movements are typically observed between 150 and 200 ms after the relevant portion of the speech input (Tanenhaus & Spivey-Knowlton, 1996), thus reflecting a normal delay between the programming of an eye-movement and its launching (Matin, Shao, & Boff, 1993).

Depending on the equipment used, the data obtained with head-mounted eyetracking offers a resolution of 33 ms at the worst, improving up to 5 ms with newer cameras (Tanenhaus & Spivey-Knowlton, 1996). By plotting the data against time, the method thus allows fine-grained, continuous monitoring of the participants' reactions to spoken input, which are assumed to reflect activation levels of words in the mental lexicon.

Indeed, Allopenna et al. (1998) have shown that using a simple but explicit linking hypothesis between activation levels and eye-movements, the output of a model of spoken word-recognition such as TRACE can be transformed into prognostics about fixation proportions which map impressively well onto fixation proportions observed in behavioral data. In particular, Allopenna et al. managed to model the fixation proportions observed in their *beaker-speaker* experiment, in which they showed that words rhyming with the target are also activated by the phonetic input: the curves they obtained by converting TRACE activation levels prognosticated that, as the input from the target word (*beaker*) unfolded over time, at first the fixations to both target and competitor (e. g. *beetle*) would rise relative to rhyming and unrelated items, then, before target and competitor begin to diverge (e. g. during the /iː/ common to all three nouns, before the /k/ in *beaker*), the rhyme (*speaker*) would begin to rise above unrelated items also, although never getting as high as the competitor, and finally, late in the input word, the rhyme would be fixated more often than the competitor, due to the fast decrease in fixations to the latter—all patterns which clearly followed those observed in the experiment with human participants.

One of the main advantages of this paradigm is that the participants' task appears fairly natural: once the cameras have been adjusted, the participants may concentrate on the job at hand and forget about the cameras, so that the data should hopefully reflect natural, every-day speech processing. No

metalinguistic judgments are required, as in lexical decision, and there is no need to interrupt the speech stream, as, for example, in gating.

However, it has been questioned whether the small number of objects in a typical visual world might not have an undesirable influence on the results obtained with the paradigm.

First of all, it is possible that the participants would develop conscious strategic reactions to the frequent presence of two objects with overlapping onsets in display, thus restricting their attention to these two objects. This objection was set aside by Allopenna et al. (1998), who showed that, although in the study mentioned above, the participants activated words rhyming with the target when instructed to move objects around in a visual world, the same rhymes were not activated when the participants heard progressively longer gates and tried to identify the target. If the participants had been consciously making use of the similarity between target and rhyme, we would have expected them to be able to apply this strategy also in the gating task.

Second, it is also conceivable that the restricted context might lead people to by-pass normal language processing mechanisms, by activating only those nouns corresponding to objects in front of them, instead of activating all nouns in the lexicon which fit the speech input.

In answer to this, Dahan, Magnuson, & Tanenhaus (2001a) and Dahan, Magnuson, Tanenhaus, & Hogan (2001b) conducted two sets of experiments which showed that eyetracking results are subject to effects from the general mental lexicon. In the first series of experiments, they presented evidence that high-frequency competitors (for instance *bed*, given target *bench*) were fixated more often than low-frequency competitors (such as *bell*). If the activated competition set had been restricted to the displayed objects, we would not have expected the overall frequency of words to have such an impact. In the second series of experiments, they additionally showed that misleading co-articulatory cues could induce participants into activating nouns which were not pictured: when the beginning of *neck* was spliced into *net* in the presence of a *net* but no *neck*, the recognition of *net* was delayed, thus reinforcing the position that the entire lexicon, including non-displayed words, participates in the results obtained with eyetracking.

1.5 Restricting the Competition Set in Native Listening

Building on the studies described in Section 1.3, Dahan and colleagues (Dahan, Swingley, Tanenhaus, & Magnuson, 2000) were also the first ones to use eyetracking in visual worlds to further explore the influence of gender on spoken word-recognition.

First of all, in two experiments with phonologically overlapping targets and competitors, they contrasted a condition in which gender information was not available prior to the noun with a condition in which the article preceding the noun carried gender information. In the first experiment, they used the plural article, which in French is gender-neutral, as in *"Cliquez sur les*$_{\text{PLURAL}}$ *boutons*$_{[masc]}$*"* (English: "Click on the buttons"), opposing this to the gender-marked singular article in the second experiment (*"Cliquez sur le*$_{\text{MASC-SG}}$ *bouton*"; English: "Click on the button"). What they found was that in the absence of gender-marking before a noun, all words which shared similar onsets with the target (for example *bouteilles*$_{[fem-pl]}$) were activated and competed for recognition, but that when gender information was available, it was used from the very start to restrict the competition set to those nouns which not only shared the same onset as the target but also agreed in gender with the preceding determiner.

This data seems to suggest that the effect of gender on spoken word-recognition takes place very early during the overall process, namely during the activation phase. Indeed, the gender information carried by the article appears to have hindered disagreeing nouns from entering the competition set, before competition took place: these nouns were not activated together with gender-matching nouns and then rejected during competition, as would have been the case if gender information had worked by selecting matching nouns after their previous activation.

Given these results, we can now risk an explanation of the processes at work behind the observations described in Section 1.3, and we can also explain what we mean by the term "gender effect": namely, a potential **facilitation or inhibition of the recognition of spoken nouns due to the reduction of the competition set after a gender-marked word**. Indeed, if listeners really do use gender information in this way to potentially narrow down the initial

set of lexical candidates during spoken word-recognition, then in the case of a congruent gender-marked article preceding a noun, recognition should proceed more rapidly and require less effort (facilitation), whereas in the case of an incongruent gender-marked article, recognition should proceed more slowly and require more effort (inhibition), because the initial competition set does not contain the correct target word, which would later have to be recovered in order for recognition to be successful.

However, in an additional set of trials in their first experiment, Dahan et al. (2000) also showed that gender could not be similarly exploited when the target noun did not overlap in onset with the other words on the screen: given target le_{MASC} *zèbre* (English: "zebra"), the gender-matching distractor $balai_{[masc]}$ (English: "broom") was not activated any more than the gender-mismatching distractor $chaussette_{[fem]}$ (English: "sock"), although it is also possible that the added activation of the gender-matching distractor was too weak or too short to be observed.

Thus, it seems that gender information carried by the article can not be used on its own to immediately pre-activate all gender-matching nouns in the lexicon (or exclude gender-mismatching nouns) before the beginning of the noun is heard; as a consequence, although in Section 1.3 we could not, on the basis of the offline studies alone, exclude spreading-activation within the lexicon, such an explanation now appears to be unlikely, because on such account, one would have expected more looks to gender-matching pictures than to gender-mismatching pictures after the article.

1.6 The Gender Effect: Form-Based or Grammar-Based?

In the previous Section, we established that the gender effect takes place during the activation phase of spoken word-recognition. Having thus clarified *when* the gender effect takes place, it is now appropriate to ask *how* such an effect might influence spoken word-recognition. Dahan et al. (2000) mention two possibilities, between which their experiments cannot distinguish: form-based or grammar-based.

In the first case, it is the surface form of the article and of the noun onset, namely a sequence of phonemes, which would influence the recognition of

the full noun, due to the increased frequency of occurrence of this particular phoneme sequence before a given noun-completion. For example, in French, after hearing the sequence /ləbu/ (*"le bou..."*), the probability that the target word should be completed as $bouton_{[masc]}$ would be higher than that of a completion by $bouteille_{[fem]}$, because people are likely to have heard the sound-sequence /ləbutõ/ (*le bouton*) any number of times previously, by comparison with */ləbutɛj/ (**le bouteille*), which they are unlikely to have ever heard before. This would imply that gender categories from the grammatical level of processing would have no impact on spoken word-recognition, although they might be used at other levels of language processing.

Although it may not generalize well to all gender languages, there are several reasons why such a reasoning would apply in particular to French and Italian, which are the languages in which gender effects in spoken-recognition have been observed (see Sections 1.3 and 1.5): in both languages, the null-article is rare (depending on the sentence, a definite or indefinite article must most of the time be introduced as a placeholder), there is no declension, and it is more common for adjectives to follow their noun than to precede them (although there are exceptions). As a consequence, articles and nouns may have a higher frequency of co-occurrence in these languages than in others.

Alternatively, the gender effect found in the studies described above might influence spoken word-recognition via grammatical information. If this were the case, it would not be the article and the noun's phonetic realizations, but instead the gender category of the article combined with the noun onset which would facilitate noun recognition by reducing competition. Thus, after hearing *"Cliquez sur le bou..."*, people would extract from *le* the information that this article was masculine, and would consequently expect *"bou..."* to be completed as a masculine noun, for example *bouton*, because of the frequency of occurrence of masculine determiners before masculine nouns.

There are various ways these two alternatives could be teased apart. Dahan et al. (2000), for their part, mention the possibility of using gender-providing contexts which have a lower frequency of co-occurrence with the target nouns. For example, one could use lower frequency determiners, such as possessives or demonstratives, or interpose various other gender-unmarked words between determiner and noun, or also use lower frequency words, such as gender-marked adjectives, to provide gender information. Indeed, if a gender effect could be at-

tested with such lower-frequency gender-marked contexts, this would speak for a grammar-based account, whereas if no gender effect was found, it would then suggest that the effect observed with high-frequency article-noun combinations had been superficial (although such a null-effect could not prove anything).

This latter solution has been employed by Dahan and colleagues, who, in an unpublished study (referred to in Dahan et al., 2000), used the same nouns as in their previous study but varied the spoken instructions: this time, gender information was carried by a low-frequency, gender-marked pre-nominal adjective (for example *astucieux*$_{\text{MASC}}$, /astysjø/, *astucieuse*$_{\text{FEM}}$, /-øz/; English: "cleverly constructed"), which also had to begin with a vowel-sound so that the preceding article would be elided, thus neutralizing gender information on the definite article: *"Cliquez sur l'astucieux bout..."*. In their results (which were, however, only preliminary) they observed that both gender-matching and gender-mismatching competitors were activated, meaning that gender was not able in this situation to exclude mismatching nouns from competition, and thus rather supporting a form-based explanation of the gender effect.

In this thesis, we pursued another course in order to make out whether the gender effect is form-based or grammar-based. We investigated the behavior of non-native listeners as far as potential gender effects on spoken word-recognition are concerned, and used the results of our observations to further research the origin of the gender effect.

Recall that in Section 1.1, we reviewed evidence that non-native listeners do not seem to deactivate their native lexicon when processing foreign language, and that on the phonetic level, native categories have a tendency to "capture" non-native input. We then wondered whether a similar "capturing" effect might also be found on the morphosyntactic level: given that we know that gender can potentially reduce the competition set in native processing, what would possibly happen in non-native processing?

Our experiments used head-mounted eyetracking in visual worlds, as described in Section 1.4. The participants were native listeners of French and German, as well as Francophone learners of German and Germanophone[4] learners of French. Both French and German have masculine and feminine nouns,[5] but although in the case of cognates, gender categories from one language frequently carry over to the other, this is not always the case: for example, *canon*

(/kanõ/; English: "canon") is masculine in French, but *Kanone* (/kanonə/) is feminine in German.

The stimulus displays were made up of a target and a competitor whose onsets overlapped phonemically across languages, accompanied by two distractors. For example, the French noun *cassette*, which in German is *Kassette* (English: "tape") was paired with *canon*/*Kanone* (English: "canon").

In Experiment 1, the instructions were in French, similar to those in the Dahan et al. (2000) study: *"Cliquez sur le/la..."*. Additionally, also as in Dahan et al., in some cases, the gender information carried by the article agreed with the competitor, whereas in other cases it did not. However, while in Dahan et al., singular and plural forms were contrasted, we used singular everywhere, contrasting instead French noun pairs with the same gender (e. g. $film_{[masc]}$/$filtre_{[masc]}$) and French noun pairs with differing genders (such as $cassette_{[fem]}$/$canon_{[masc]}$). In spite of this difference, we anticipated to replicate the findings of Dahan et al. (2000) with Francophone native participants, by finding competition in the case of same-gender noun pairs and no competition in the case of different-gender noun pairs.

The same experiment was also run with a group of proficient Germanophone learners of French, in order to examine how non-native listeners would react to French gender. Indeed, a crucial point of the setup was that even when the noun pairs did not match in gender in French, they always matched in gender in German (as in $Kassette_{[fem]}$/$Kanone_{[fem]}$), because the competitors were chosen so that the French noun and its German translation would be of different genders, as in the *canon*/*Kanone* example above. Thus, although when target and competitor noun matched in gender in both French and German, we reckoned with the same pattern as in the Francophone natives, the results when target and competitor noun differed in gender in French (but not in German)

[4]Among our French-speaking participants, we had some French-speaking Belgians, and among our German-speaking participants there was one German-speaking Swiss. Thus we refer to the French-speaking participants as Francophones (instead of French) and to the German-speaking participants as Germanophones (instead of German).

[5]The French gender system being limited to masculine and feminine, we used no neuter nouns in German, in order to allow a direct comparison between the two languages, although it seems that the German neuter actually falls together with masculine in French, as confirmed by Corbett (1991, p. 316), who says that, in French, neuter and masculine have combined over the course of history.

should allow us to observe whether the non-native listeners were able to use French gender in a native fashion, behaving similarly to the Francophones, or whether they would be incapable to master French gender, showing competition where the Francophone natives did not.

Assuming then that the Germanophone non-native listeners were unable to put French gender to use, thus showing a competition effect with French noun pairs of opposing genders, two possibilities would then remain: did the non-native participants use no gender information at all, or did they draw on gender information from their mother-tongue? In both cases, we would expect to see competition: if the participants were not using gender, the onset overlap of target and competitor should lead to competition, just as if they used gender information from the mother-tongue, the absence of a gender difference between target and competitor in German should lead to competition, because gender could not be used to exclude the competitor from the competition set.

This was investigated in Experiment 2, in which the same materials were used, but the instructions were in German. As in French, the instructions contained a definite, gender-marked article providing gender information immediately before the noun. However, since in Experiment 1, the German competitors (e. g. $Kanone_{[fem]}$) always had the same gender as the target ($Kassette_{[fem]}$), in this experiment, there should be no reason why gender should help reduce competition, if we considered only the language spoken during the experiment. Thus, in the case of Germanophone native participants we reckoned with competition in both the French same-gender pairs and different-gender pairs.

More interesting was running the experiment with Francophone learners of German. Here, we envisaged two possible patterns of behavior: if the non-native participants made no use of gender information at all, we expected competition with both same-gender and different-gender pairs, due to the onset overlap of target and competitor, but if they relied instead on gender information from their mother-tongue, it was possible that, in the case of different-gender pairs, they would not activate the competitor (for example *Kanone*), because its French translation ($canon_{[masc]}$) did not have the same gender as that of the target and thus did not agree in gender with the the article in the spoken instruction (for example, *"die*$_{\text{ART-FEM}}$ *Ka. . . "*). This latter case would then represent a case of morphosyntactic "capturing", since words whose gender is different in the mother-tongue and the foreign-language (for example $canon_{[masc]}/Kanone_{[fem]}$)

are "captured" by the native category (here masculine), and grouped into the wrong gender category.

It is at this point that we now come back to the debate on the origin of the gender effect, for depending on our results, we would also be in a position to assert whether the gender effect observed is form-based or grammar-based.

Indeed, the surface form of the various definite articles is very different in French and German: *die* (/diː/, German feminine) is unlike *la* (/la/, French feminine), and *der* (/deɐ̯/, German masculine) is unlike *le* (/lə/, French masculine). Consequently, if the gender effect were form-based, non-native listeners would have to re-learn the use of gender all over again for each language. Two possibilities would then have to be distinguished. If they were successful, we would expect the non-native participants in both experiments to behave like the respective groups of native participants: in particular, the Germanophone learners of French should then use French gender in the different-gender pairs of Experiment 1b to reduce competition, while the Francophone learners of German should not reduce competition in the different-gender pairs of Experiment 2b. On the other hand, if non-native listeners never mastered gender in their foreign language, we would expect to see competition in both conditions (same-gender and different-gender pairs) of both experiments, because gender had no influence at all on the process of spoken word-recognition, the competition set being determined only by the onset overlap of words in the mental lexicon.

If the gender effect were grammar-based, the same two cases would also have to be distinguished. If non-native listeners were able to learn to use the categories of their foreign language, we would, once again, expect the same behavior as natives. On the contrary, if non-native listeners revealed an incapacity to use foreign gender categories, then either the Francophone learners of German would use the categories of their mother-tongue to wrongly suppress competition in the different-gender pairs of Experiment 2b, or, if the non-native participants ignored gender altogether, we would, as above, observe competition with both groups of non-natives in all conditions, because of the lack of influence from gender.

2 Experiment 1: French Instructions

Method

Participants

Experiment 1a: Native Listeners. Thirteen Francophones took part in Experiment 1a. All had normal or corrected-to-normal vision and normal hearing. They were all offered 5 euros for taking part in the experiment, although some acquaintances of the author refused the money and took part out of friendship.

Among the participants, nine were born in France, three were Francophone Belgians, and one was born in Germany as the daughter of a French military. They were on average 33 years old, and had studied German on average for 3.77 years in school or college, but all admitted on their own account that they could not speak it. On average, they had spent a total of 10 months in Germanophone countries in the course of their lives, but only one had been exposed to German as a child, the daughter of a French military stationed in Germany, who had gone to a German kindergarten, but as an adult could not speak German anymore.

Experiment 1b: Non-Native Listeners. Twenty-three Germanophones who had learned French as a second language participated in Experiment 1b. As in Experiment 1a, all had normal or corrected-to-normal vision and normal hearing, and all were offered 5 euros for participating, though some did not take the money.

All participants were from Germany, except one who was from the German-speaking part of Switzerland. The mean age of the participants was 30. On average, they had studied French for 7.43 years in school, starting at the mean age of 13. In all cases, their own personal appreciation of their level of French

was that they could speak it well, and they felt confident enough to participate in an experiment which was run entirely in French.

Materials

Same-Gender Vs Different-Gender Pairs. Target and competitor nouns were selected using a French-German dictionary (Sachs & Villatte, 1968). The constraints at this point were that the French nouns and their German translations overlap at onset by at least two phonemes (e. g. the French word for "tape" and its German translation, respectively *cassette* and *Kassette*, which share /ka/ as their onset), that the German nouns be either masculine or feminine (no neuter),[5] and that all nouns refer to picturable objects.

From this list of nouns, 35 preliminary target-competitor pairs were built by associating nouns which overlapped at least by two phonemes, both in French and in German: the French target *cassette* overlaps with the competitor *canon* (English: "canon"), while at the same time in German *Kassette* overlaps with *Kanone*.

Among these 35 pairs, there were two kinds of items: although in all cases, the target was of the same gender in French and in German (French *cassette* and German *Kassette*, for example, are both feminine, whereas French *film* and German *Film* are both masculine), the gender of the competitor divided the items in two categories. In the first group, which comprised 18 target-competitor pairs, the competitor was of different gender in the two languages, with the German translation being of the same gender as the target (*Kanone* is feminine) and the French noun being of the gender opposing the target and its own German translation (*canon* is masculine; see Figure 1a). These pairs of nouns will henceforth be referred to as "different-gender pairs", since in the tested language (French), target and competitor differed in gender. In the second group (17 noun pairs), on the other hand, the competitor was of the same gender as the target in both languages (the French noun *filtre*, which was the competitor for the masculine target *film*, and its German translation *Filter*, are both masculine; see Figure 1b). These items will henceforth be termed "same-gender pairs", since in language of the experiment, target and competitor shared the same gender.

Different-Gender:

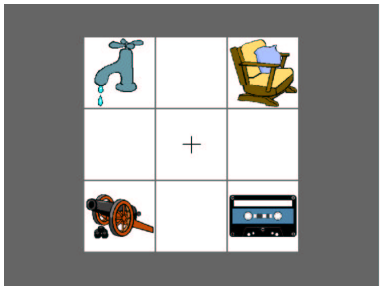
	Target	Competitor
French	cassette (f)	canon (m)
German	Kassette (f)	Kanone (f)

(a)

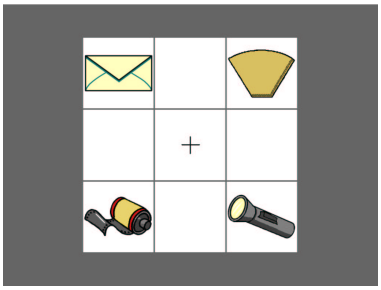
Same-Gender:

	Target	Competitor
French	film (m)	filtre (m)
German	Film (m)	Filter (m)

(b)



(c)



(d)

Figure 1: Two target-competitor pairs, one from different-gender condition (Figure 1a) and one from the the same-gender condition (Figure 1b), displayed together with distractors above their corresponding grids (Figures 1c and 1d), such as these appeared on the screen during the experiment. The letter "m" stands for masculine and "f" for feminine.

Pretest. All pictures were taken from the IMSI MasterClips Collection (1990) Image Collection, except for a few which were found on the world wide web. They were simple colored drawings such as might appear in children's picture books. In order to ensure that the pictures would be recognized for what they were by the participants during the experiment, written questionnaires were used to pretest them: 10 Francophone and 10 Germanophone native speakers were asked to name the pictures in their mother-tongue, and 10 more Germanophone native speakers were asked to rate how easily recognizable the pictures were.

In the French and German naming tests, a total of 86 pictures were presented to the participants (both nouns from the 35 target-competitor pairs, plus 16 backup pictures to be used as replacements for those target or competitor pictures which did not work well), whereas in the rating test, six pictures were tested in two different versions, so that overall, there were 92 pictures to be rated.

There were eight pictures per page, and the same randomization was used for all questionnaires, but in half of the questionnaires, the order of the pictures

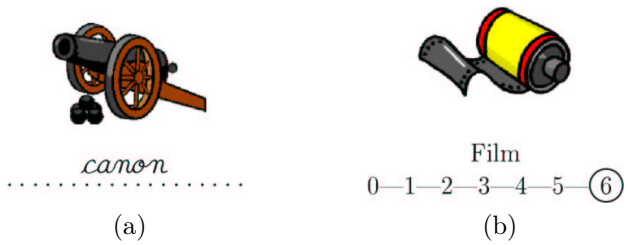


Figure 2: Two pictures such as they appeared in the French naming pretest (Figure 2a) and in the rating pretest (Figure 2b). In the naming pretest, the dotted line was left blank, and we expected the participants to name the object a "canon".

was reversed. The instructions and an example appeared on the first page. For the naming test, a dotted line appeared under each picture (as in Figure 2a), whereas for the rating test, the expected name of the picture appeared above a seven-point scale ranging from zero (difficult to identify) to six (easy to identify; see Figure 2b).

The answers to one French and one German naming questionnaire were discarded, and the questionnaires replaced, the first because the participant left a great number of blanks (18 from 92 pictures), and the second because the participant often did not name the object represented by the picture, but instead reacted with semantic associations in 46 cases (writing, for example, *Leiche*, which is German for "corpse", underneath the picture of the empty car trunk).

In the evaluation of the naming test, we accepted a number of "good" synonyms as correct answers, since Zwitserlood (1989) has shown that synonyms are activated during spoken word-recognition, although in some cases, they were not of the same gender as the original noun: whereas *antilope* and *gazelle* happen to be both feminine in French and German, the French word *pellicule*, which is a common synonym for the masculine *film*, is itself feminine.[6]

Final List of Items. Based on the results of the pretest, we then excluded some nouns for which we could not produce easily recognizable pictures (pictures with an average rating score of less than 2.5, or pictures which were often

[6] Another problem also arises with German compound words, which are always of the gender of the last compound. Thus although *Tonkassette* (English: "audio cassette") is feminine just as *Kassette* (but unfortunately differs in onset), *Filtertüte* (English: "paper filter"), unlike *Filter*, is feminine, because its last compound, *Tüte* (literally "paper bag"), is feminine.

given names other than those we had expected, such as our *villa*, which was often termed a "house", in French *maison* and in German *Haus*), improved some pictures whose contours or colors were too pale, according to the suggestions of the pretest participants, and then put together a final list of 30 target-competitor pairs (see Appendix, p. 79): 15 "same-gender pairs" and 15 "different-gender pairs".

In this final list, the agreement between the names we intended and the participants responses in the pretest was 87.3 % in French and 87.7 % in German, and the recognizability of the pictures was rated with a mean score of 4.7. Many nouns were French-German cognates.[7] The French targets and competitors all overlapped at onset by a minimum of two phonemes; five noun pairs had an overlap of three phonemes (based on the phonemic transcriptions in Le Robert, 1995, and Duden, 1990). The overlapping sounds were either a consonant and a vowel or, in five cases, a consonant and a trill or liquid (/r/ or /l/). The onset overlap between the French nouns and their German translations, between the German translations of the targets and those of the competitors, and between the French targets and the German competitors was similar. Most pictures were inanimate objects, except for two animate targets and two animate competitors in the different-gender pairs, and two animate targets and one competitor, as well as two "place names" (i. e. slightly more complex pictures depicting not an object, but a landscape) in the same-gender pairs.

Moreover, since Allopenna et al. (1998) have shown that words that rhyme with the target are also activated during the spoken word-recognition process, we avoided rhyming targets and competitors, and also took care to avoid semantically related nouns, of which studies have shown that they may activate one another (Zwitserlood, 1989).

Finally, given that Dahan et al. (2001a) have shown that the probability of fixating a competitor that matches the phonemic onset of the target noun varies with its lexical frequency, we also made sure that the overall frequency of the targets and competitors did not differ reliably within any of the two conditions, using the Brulex database (Content, Mousty, & Radeau, 1990) for French and

[7] Cognate: Two words in related languages are cognate if they come from the same original word. Generally cognates will have similar, though often not identical, sounds and meanings (Bickford & Tuggy, 2002).

webcounts from AltaVista (`www.altavista.com`) for both languages.[8] Using the webcount figures, the difference between the mean frequency of the targets and the mean frequency of the competitors was not statistically significant in any of both conditions (results of the ANOVA for the same-gender pairs in French: $F[1, 28] = 0.374$, $p = 0.546$; for the different-gender pairs in French: $F[1, 28] = 2.321$, $p = 0.139$; for the same-gender pairs in German: $F[1, 28] = 0.681$, $p = 0.416$; for the different-gender pairs in German: $F[1, 28] = 1.003$, $p = 0.325$). Using the Brulex database however, the analysis also indicated no difference for the same-gender pairs ($F[1, 28] = 1.772$, $p = 0.194$), but it did show that the competitors were more frequent than the targets in the different-gender pairs, an effect which should actually work against our predictions (mean for the targets: 600.27; mean for the competitors: 2322.6; $F[1, 28] = 7.812$, $p = .009$).

Note that in the same-gender pairs, we could also have alternated the role of the targets and competitors in the experiment, given that the gender of all four nouns, target and competitor in both languages, was the same (e. g. French *film* and *filtre*, as well as German *Film* and *Filter*, are all masculine), in order to avoid a possible bias due to frequency or to the quality of the pictures (as was done by Dahan et al., 2000). However, this was not possible for the different-gender pairs, because the French competitor was of the gender opposing the three other nouns (*canon* is masculine, whereas *cassette*, *Kassette* and *Kanone* are all feminine), so that we chose instead not to alternate the nouns in the same-gender pairs, but to balance them as well as possible between targets and competitors based on frequency and picture quality.

[8] Blair, Urland, & Ma (2002), as well as Keller, Lapata, & Ourioupina (2002), have shown that, if search engines with sufficiently large databases are queried, webcounts are highly consistent with more traditional methods of estimating word frequency, such as the analysis of a traditional linguistic corpus or a lookup in a linguistic database, thus offering a potential solution to data sparseness. Being confronted with the fact that some of the nouns used in our study (for instance *Domino* or *Papaya*) were not to be found in the CELEX database for German (Baayen, Piepenbrock, & van Rijn, 1993), we thus resorted to AltaVista webcounts. This particular search engine was chosen because it has a large database (database size: 1,689 millions, according to Search Engine Showdown, 2002), because it had produced good results in both the study by Blair et al. (2002) and the study by Keller et al. (2002), and also because the retrieved webcounts were not subject to too much variation over a period of a few days, as was unfortunately the case with Google (`www.google.com`).

Distractors and Fillers. Given this final list of targets and competitors, we then added two unrelated distractors to each target-competitor pair, in order to have four pictures to display on each screen. The two distractor nouns did not overlap in French and German, and they also overlapped neither with the target, nor with the competitor, nor with each other: for example, the distractors associated with the above mentioned pair *cassette*/*Kassette* and *canon*/*Kanone* were *fauteuil*/*Sessel* (/fotœj/ and /zɛsl̩/; English: "armchair") and *robinet*/*Wasserhahn* (/rɔbinɛ/ and /vasɐhaːn/; English: "faucet").

Among the distractors, seven were animate nouns and two were "place names". We also took care to avoid having two pictures on a trial which belonged to the same semantic field. Frequency of distractors was not controlled for because Dahan et al. (2001a) have shown that when the onset of a distractor does not overlap with the target, the probability of fixating it does not vary with lexical frequency.

Additionally, we also had 33 filler-items. Our main concern here in our choice of the nouns was to minimize the possibility of participants developing expectations as to what the target would be in the course of the experiment. Overall, there were 16 animate nouns and 8 "place names" in the fillers. The gender of the fillers was basically random: in French, 76 were masculine and 56 feminine, and German, 60 masculine in German, 50 feminine, and 22 neuter.

Among the fillers items, there were none in which the target noun overlapped with another noun in the display, whether in French or in German. There were 29 items in which the French target noun did not overlap with its German translation, as opposed to an overall total of 34 items in the experiment in which the target did overlap across languages (30 experimental stimuli and 4 fillers). In 17 fillers, none of the four nouns overlapped with another noun, whether across languages or between words, as in example 1 in Table 1; in eight, two nouns, one of which the target, did not overlap with the other nouns in any way, but the two other nouns overlapped with one another and also overlapped with their own translation, as in example 2; in four, two nouns, including the target, did not overlap across languages, but the two other did, with no between-word overlap, as in example 3; and in the remaining four, all four nouns overlapped with their own translation, also with no between-word overlap, as in example 4.

		Targets				Count
1	F	jambon	moufle	cadre	puits	17
	G	Schinken	Handschuh	Rahmen	Brunnen	
	E	"ham"	"mitten"	"frame"	"well"	
2	F	calculatrice	éponge	pagode	parcmètre	8
	G	Taschenrechner	Schwamm	Pagode	Parkuhr	
	E	"calculator"	"sponge"	"pagoda"	"parking meter"	
3	F	sèche-cheveux	brouette	rose	planète	4
	G	Fön	Schubkarre	Rose	Planet	
	E	"hair drier"	"wheelbarrow"	"rose"	"planet"	
4	F	violon	balai	flamme	robot	4
	G	Violine	Besen	Flamme	Roboter	
	E	"violin"	"broom"	"flame"	"robot	

Table 1: Examples of filler-items. The letters "F", "G" and "E" on the left-hand side stand for French, German, and English (translations).

It was hoped that, due to the presence of many words which overlapped across languages (including many cognates) but were not targets (36 nouns distributed over the last three categories of fillers), we might avoid participants developing a strategy such as *"when the onset of a noun on the grid is the same in both languages (*or: *when they are cognates), then it is likely to be the target"*. Furthermore, the presence of 8 fillers in which two nouns overlapped without the target being one of them should also help to avoid that the participants would come up with the expectation that *"when the onset of two words on the grid overlap, then the target is one of them"*.

Experimental Displays. The grid was made of black lines delimiting nine white squares on a dark gray background, and had a small black cross in the middle square. It was designed so that each square would occupy about one fourth (191 pixels, corresponding to ca. 7.5 cm) of the height of the screen, which had a resolution of 1024 × 768 pixels. The pictures were scaled to a maximum height and width of 185 pixels, so that each would fit in one square of the grid without touching the lines, but still making full use of the available space (example grids can be seen in Figures 1d and 1c).

The four pictures were each assigned randomly to one of the corner squares of the grid, with the additional constraint that, for the experimental trials, targets and competitors should occur equally often in each condition in the upper and lower half of the grid. Finally, for the experimental trials also, the color of the pictures belonging together was slightly adapted so that none of the pictures would stand out too much.

Spoken Instructions. The target nouns were embedded in French carrier sentences, which instructed the participants to click on one of the pictures displayed on the screen. An instruction was for instance *"Cliquez sur la cassette"* (English: "Click on the tape"), the gender of the upcoming noun being foretold in the instruction by the use of the congruent article: *la* (/la/) for feminine nouns, and *le* (/lə/) for masculine ones (e. g. *"Cliquez sur le film"*). Only the target noun was named in the instructions, the competitors and distractors being left completely unmentioned, so that the participants did not at any point during the experiment hear the competitors preceded by the correct gender-congruent article.

Note that although Guillelmon & Grosjean (2001) mention that it is possible that English natives having learned French relatively late may not be able to discriminate the feminine (*la*, /la/) and masculine (*le*, /lə/) articles in French, it is unlikely that this should be the case with our Germanophone natives who had learned French as a second language, because German, just as French, distinguishes the phonemes /a/ and /ə/.

The spoken instructions were recorded onto a digital audio tape by a Francophone native speaker (the author) in a sound proof room, and then transferred to disk by playing them back using a DAT-recorder and recording them onto the computer. Each instruction was then cut with `Praat` (Boersma & Weenink, 2003) and stored as a separate file. A blank of 500 ms was added at the beginning, and the onset time of each target noun and duration of the overlap were measured on the basis of the recordings: the mean noun onset time from the beginning of the instruction was 632 ms, and the mean duration of the putative overlap (from noun onset to disambiguation point in the target noun, e. g. the duration of the /ka/ when *cassette* is opposed to *canon*) was 139 ms (mean for the same-gender pairs: 128 ms, mean for the different-gender pairs: 150 ms).

Procedure

Instructions. The experiment was conducted entirely in French with all participants, whether they were native speakers or whether they had learned French as a second language. At the beginning of the experiment, the participants were given written instructions (also in French) containing an explanation of the task, telling them that they would see displays made up of a grid with four objects on it, and that a recorded voice would ask them to click on one of the objects.

They were told that their reaction times did not matter. An example of a trial display and an accompanying directive to click on one of the objects were printed on the instruction sheet.

After giving the participants time to read the written instructions and answering any remaining questions concerning the procedure, we helped them determine which was their dominant eye. They were then seated in front of the middle of a 19-inch monitor, approximately sixty centimeters away, and were given the headphones.

Eyetracking. The participants' eye-movements were monitored using an SMI head-mounted eyetracker with EyeLink software, supplemented by some custom developed software. Although the SMI eyetracker has two lightweight cameras mounted on its headband, one for each eye, we only used the one corresponding to the participant's dominant eye, the other camera being bent back over the participant's head during the experiment.

The center of the pupil was tracked in order to determine the position of the eye relative to the head, and at the same time, the position of the head relative to the screen was also monitored. From this, throughout the experiment, the software computed what point on the screen the participants were looking at, and recorded the onset and offset times and spatial coordinates of the participants' saccadic eye-movements and fixations, sampling every four milliseconds. Each picture measured approximately 7.5×7.5 centimeters, thus corresponding to a visual angle of $7\,^{\circ}$, which is well within the resolution of the eyetracker ($0.1\,^{\circ}$).

Additionally, the coordinates of the participant's mouse clicks were also recorded, and our custom software, knowing which object was the target, determined whether the participant had clicked on the correct object or not, rewarding them with a high-pitched chime if they were right but playing a low-pitched bong-sound if they were wrong.

After the eyetracker was calibrated, the experiment began. First, the participants were shown three practice items, taken from the fillers described on page 42. Then, after the participants had had a chance to ask any further questions, the actual experiment started with a filler, which was followed by the first experimental stimulus. Filler and experimental trials then alternated

throughout the experiment, the same random order being used forwards for one half of the participants and backwards for the other half.

The sound, which included a 500 ms blank at the beginning, started playing 50 milliseconds after the display of the grid, so that the spoken instruction started approximately 550 milliseconds after the participants saw the objects on the grid. This was similar to the procedure used by Dahan et al. (2000), who, in comparison to Allopenna et al. (1998), reduced the time between appearance of the grid and the start of the instruction from 2 seconds to 500 milliseconds, in order to avoid the participants' consciously naming the pictures on the screen and thus retrieving their gender.

Moreover, also as in Dahan et al. (2000), there was no explicit instruction to fixate the cross at the onset of each trial, thus making the setup more natural for the participants, but with the disadvantage that participants could at the beginning of a trial be fixating any object on the grid.

Between trials, a centered fixation point appeared on the screen, and participants were instructed to look at it. The experimenter could then correct potential drifts in the calibration of the eyetracker. There was a possibility of a pause after 30 trials, after which the experiment went on again with a filler.

Language Background. After the experiment, the participants were asked to fill in a written questionnaire giving us some information about their linguistic background. The questionnaire was also completely in French. First, the participants were asked some general questions such as in what countries or linguistic regions they had ever lived for more than six months, what dialects might possibly have been spoken at home during their childhood, what languages they could understand easily, which language they considered as their mother-tongue and what foreign languages they studied in school and for how long. Finally, the non-native participants were given a list of all target and competitor nouns in the experiment, and were asked to circle the correct article (*le* or *la*) for each of them, thus revealing whether they knew the correct gender of the noun or not. Altogether, the experiment and the completion of the questionnaire took approximately half an hour.

Analysis. The data was coded using graphical software which superimposed the participants fixations on template-grids similar to those used during the

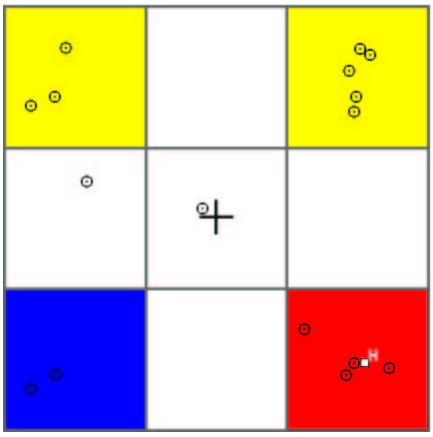


Figure 3: The template-grid which was used to code one participant's fixations for the *cassette* trial. The red square indicates the location of the target picture, the blue square the location of the competitor, and the yellow squares that of the distractors. The circled dots show the participant's fixations, and the filled square next to the "H" indicates where the participant clicked with the mouse.

experiment, but where the pictures of the objects were replaced with colors indicating whether the object was a target, a competitor or a distractor (see example template in Figure 3). Fixations on the lines of the grid, on the cross, on the white squares or on the gray background were all coded as "other". Only fixations with a minimum duration of 80 milliseconds were taken into account, blinks being added to the previous fixation and saccades discarded.

Results

Experiment 1a: Francophones Listening to French

The data collected from the last participant was removed in order to retain an even number of participants, so that, overall, the forward and backward orderings of the trials were shown equally often to the participants. We then established for every item in 10 millisecond time slices what type of picture (e. g. target, competitor, distractor or "other") each participant had been fixating, from target noun onset up to 1500 milliseconds later. From this, we computed the averaged proportions of fixations to each type of picture for each condition, by participants and by items, in 10 millisecond slices, averaging the fixation proportions to both distractors.

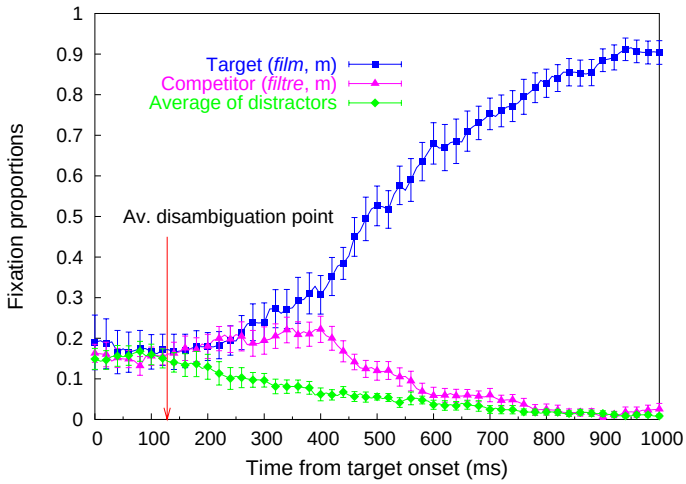
		Targets	Competitors	Av. Distractors
0–200 ms	same-gender pairs	17.4 %	16 %	14.8 %
	different-gender pairs	16.5 %	11.9 %	14.9 %
200–600 ms	same-gender pairs	38 %	16.4 %	7.2 %
	different-gender pairs	36 %	13 %	9.6 %

Table 2: Average proportions of fixations for each type of picture (targets, competitors and averaged distractors) in Experiment 1a, over the time-windows from 0 to 200 ms and from 200 to 600 ms, in both conditions. The proportions do not sum up to one because the fixations to the white squares, to the gray background and to the cross and grid lines are not mentioned.

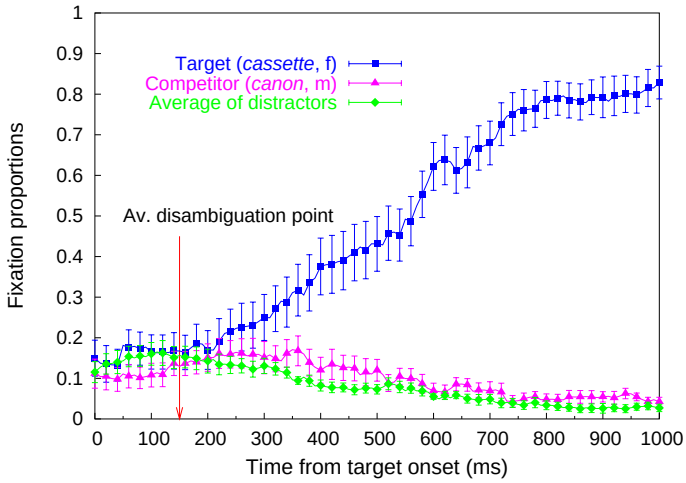
Because it is estimated that a saccadic eye-movement is typically programmed about 150–200 ms before it is launched (Matin et al., 1993), it is expected that the acoustical input will begin to influence the observed fixations starting approximately 200 ms after the playing of the corresponding part of the spoken instructions; this was confirmed for lexical access by Allopenna et al. (1998), who observed an increase in fixations to target and competitor nouns 200 ms after target onset. Consequently, in our experiment, any difference in the proportions of fixations before 200 ms would have to be attributed to noise or to a bias due to the pictures themselves, because the participants found one picture type (e. g. the competitors) more interesting to look at than another (e. g. the distractors).

Figure 4 on page 50 presents the averaged proportions of fixations by participants, from 0 to 1000 ms after target onset, first for the same-gender pairs (Figure 4a) and then for the different-gender pairs (Figure 4b). As can be seen in Figure 4a, in the case of the same-gender pairs, the proportion of fixations to competitors and distractors begin to diverge about 200 ms after the onset of the target noun, the proportion of fixations to the competitors remaining higher until approximately 500 ms after target onset. By comparison, Figure 4b shows that in the different-gender trials, the proportion of fixations to the competitor and distractor did not differ much over time: although the curves do slightly diverge between 200 and 500 ms, the overlap of the standard error bars indicates that the difference was probably not significant.[9]

[9]This was confirmed by an ANOVA on the mean proportions of fixations from 200 to 500 ms: $F_1[1, 22] = 2.476$, $p = 0.144$; $F_2[1, 28] = 3.783$, $p = 0.072$.



(a) Same-gender pairs



(b) Different-gender pairs

Figure 4: Averaged proportions of fixations by participants in Experiment 1a, from 0 to 1000 ms after target onset. The vertical bars show the standard error.

In order to statistically compare the proportion of fixations to each picture, we defined a time-window extending from 200 ms to 600 ms after target onset, over which the fixations to each type of picture were averaged, for each participant and each item, in both conditions (same-gender pairs and different-gender pairs). The average proportions of fixations to each type of picture over this time-window are shown in Table 2. Similarly, the fixations to each type of picture were also averaged over the time-frame extending from 0 ms to 200 ms after target onset, so as to be able to test that any difference we found between the proportion of fixations to the competitors and to the distractors could be ascribed to language processing and not to a possible picture bias.

We conducted one-factor ANOVAs on the mean proportions of fixations over the above-mentioned time-windows, with picture type (competitor or distractor) as within-factor. Over the time-window from 0 to 200 ms, the analysis showed that there was no difference in any of the two conditions between the proportion of fixations to the competitors and distractors (same-gender pairs $F_1[1, 22] < 1$, $p = 0.591$; $F_2[1, 28] < 1$, $p = 0.711$; different-gender pairs: $F_1[1, 22] < 1$, $p = 0.343$; $F_2[1, 28] < 1$, $p = 0.435$). This lack of picture bias in the initial time-window consequently justified a direct comparison of the further course of the fixation proportions.

We then ran similar ANOVAs over the time-window from 200 to 600 ms after target onset. In the case of the same-gender pairs, we expected to find what is called a "competition effect", the competitor being fixated more than the distractor by virtue of its onset overlap with the target noun. This would suggest that upon hearing the first few sounds of the target noun, during the time when competitor and target putatively overlapped (before the disambiguation point), both target and competitor are activated in the mental lexicon, by comparison with unrelated nouns such as our distractors. Indeed, in this condition, the article preceding the noun in the spoken instructions agreed with both the French target and competitor (e. g., in *"Cliquez sur le fil... [m]"*, the article, le_{MASC}, agrees with both $film_{[masc]}$ and $filtre_{[masc]}$), so that before the disambiguation point, there was no way for the participants to know which noun they would be asked to click on. The expectation of a competition effect was confirmed by the results of the ANOVA, which showed that the Francophone listeners fixated the competitor more often than the distractor ($F_1[1, 22] = 30.711$, $p = 0.000$; $F_2[1, 28] = 7.943$, $p = 0.014$).

On the contrary, in the case of the different-gender pairs, the French article agreed with the target only and not with the competitor (e. g., la_{FEM} in *"Cliquez sur la ca. . . [ssette]"* agrees with $cassette_{[fem]}$, but not with $canon_{[masc]}$), so that it was theoretically possible that, using gender information, the participants might guess which would be the target by hearing only the first few sounds of the word, in which case we should not observe competition. However, if the participants did not make use of the French gender information available from the article, then no choice of alternative between target and competitor would be possible until after the disambiguation point had been reached. Here, the statistical analysis lead us to favor the first possibility, because the the statistical analysis showed no significant difference between the fixations to the competitors and to the distractors ($F_1[1, 22] = 3.180$, $p = 0.102$; $F_2[1, 28] = 3.002$, $p = 0.105$), thereby revealing no initial competition of the target and competitor nouns during lexical access.

This suggests that, in the different-gender pairs, the Francophone native participants were able to suppress competition on the basis of the gender information carried by the article preceding the noun, and this even though the somewhat longer putative overlap of the target and competitor nouns (mean for the same-gender pairs: 128 ms; for the different-gender pairs: 150 ms) should, if anything, have augmented the competition effect in this condition.

Taken together, these results show that, when the target and competitor nouns are of the same gender, a competition effect was observed, but when the nouns are of differing gender, no competition took place, admittedly because participants could use gender information early on to disambiguate between both nouns with similar onsets.

This is very similar to what was shown by Dahan et al. (2000): where we used target and competitor nouns of the same gender to show a competition effect in French, Dahan et al. (2000) in their first experiment used the plural form of the nouns and preceding articles, which in French do not overtly mark gender. Just as we did in the same-gender condition, they also found that in the absence of overt gender-marking on the article, the competitors were fixated more often than the distractors, thus indicating that the onset of the target noun temporarily activated other overlapping items in the mental lexicon.

In order to confirm the similarity between our results and those of Dahan et al. (2000), we repeated the statistical analysis of our same-gender noun pairs

on the time-frame used by Dahan et al., which was 300 ms to 700 ms after target noun onset, and indeed, obtained corroboration: the result of a one-factor ANOVA on this time-frame was significant ($F_1[1, 22] = 34.672$, $p = 0.000$; $F_2[1, 28] = 8.336$, $p = 0.012$), the same test being non-significant between 0 and 300 ms ($F_1[1, 22] = 2.673$, $p = 0.130$; $F_2[1, 28] = 1.438$, $p = 0.250$).

Moreover, in their second experiment, Dahan et al. (2000) went on to show that when gender can help to determine the target before the point where the target noun itself is unambiguous, then gender information preceding the noun is used during lexical access, thus reducing competition. Here too, we repeated our statistical analysis of our different-gender pairs for the time-frame chosen by Dahan et al., and obtained confirming evidence: the ANOVA was non-significant ($F_1[1, 22] < 1$, $p = 0.698$; $F_2[1, 28] < 1$, $p = 0.744$), as well as non-significant over the first 300 milliseconds ($F_1[1, 22] < 1$, $p = 0.698$; $F_2[1, 28] < 1$, $p = 0.744$).

Consequently, it can be said that our results replicate those obtained by Dahan et al. (2000): in this experiment, just as in the Dahan et al. study, when gender could not be used early on to distinguish between target and competitor because the article agreed with both, competition was observed as in English, a language in which inanimate nouns do not have gender (see, for example, Tanenhaus et al., 1995, and Allopenna et al., 1998); however, when the preceding article in the instruction agreed with only one of both nouns, then gender information was used to reduce competition, thus establishing that the initial set of lexical candidates considered for during recognition of a noun can be constrained by morphosyntactic information carried by the preceding context.

Moreover, this ensures that, as we had hoped, under the appropriate linguistic conditions, the pictures we used as materials *do* in fact allow the observation of potential gender effects: the competitor pictures in the different-gender pairs were not so intrinsically interesting to look at that people would not care about gender information and still look at them even if gender agreement was violated. This thus establishes a baseline for the following experiment, in which we set out to research whether non-natives would behave similarly to the native listeners: whatever our observations, the results we obtained would in any case have to be attributed to the difference between native and non-native processing, and not to possible shortcomings of our materials.

		Targets	Competitors	Av. Distractors
0–200 ms	same-gender pairs	24.8 %	18.5 %	15 %
	different-gender pairs	24.2 %	19.5 %	14 %
200–600 ms	same-gender pairs	36.4 %	21.8 %	11.2 %
	different-gender pairs	37.2 %	19.5 %	11.9 %

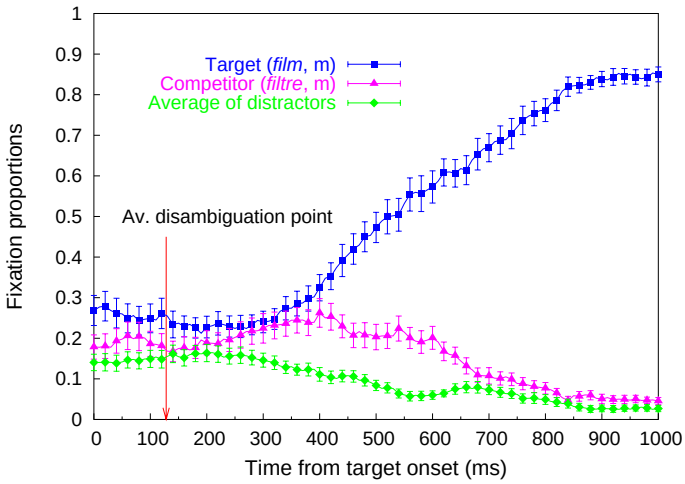
Table 3: Average proportions of fixations for each type of picture (targets, competitors and averaged distractors) in Experiment 1b, over the time-windows from 0 to 200 ms and from 200 to 600 ms, in both conditions.

Experiment 1b: Germanophones Listening to French

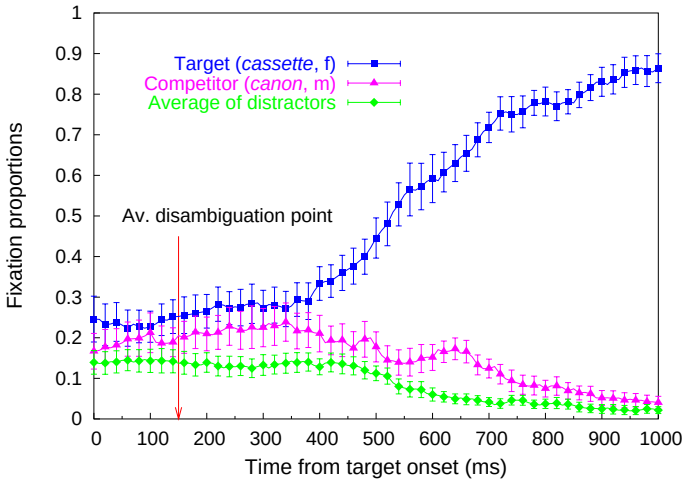
The data from two participants was removed before the analysis because, in many cases, they did not look directly at the pictures when scanning the experimental displays, but instead looked mostly at the cross in the middle of the grid and still managed to click on the correct picture (peripheral vision), a behavior which is occasionally observed in participants. Moreover, the data from the 23rd participant was removed to ensure that the forward and backward orderings of the trials were shown to the participants an equal number of times.

We also removed the fixation data from 7 trials (1.2 %) in which no sound was played due to a technical failure, as well as from those few trials in which the Germanophone non-native speakers had clicked on the wrong object (2 %), thus revealing that they did not know the French nouns used in that trial. Such a low percentage of errors suggests that the Germanophone non-native participants had no difficulty performing the task. This left us with fixation data from 291 same-gender trials and 290 different-gender trials.

Figure 5 on page 55 presents the averaged proportions of fixations for the Germanophone participants, from 0 to 1000 ms after target onset, first for the "same-gender pairs" (Figure 5a) and then for the "different-gender pairs" (Figure 5b). Again, fixation proportions for the two unrelated distractors were averaged. Both figures show a competition effect: in Figure 5a, the proportions of fixations to the competitors and to the distractors begin to diverge shortly after 200 ms, the difference probably becoming significant around 300 ms (as suggested by the lack of overlap of the standard error bars) and the curve for the competitors remaining higher until shortly before 700 ms after target onset, whereas in Figure 5b, the lines diverge somewhere between 200 and 300 ms and remain apart until after 700 ms.



(a) Same-gender pairs



(b) Different-gender pairs

Figure 5: Averaged proportions of fixations by participants in Experiment 1b, from 0 to 1000 ms after target onset. The vertical bars show the standard error.

As in Experiment 1a, we conducted one-factor ANOVAs on the mean proportions of fixations over the time-frames extending from 0 to 200 ms and from 200 to 600 ms, with picture type (competitor or distractor) as within-factor. The average proportions of fixations to each picture type over both time-windows are shown in Table 3.

Again, over the first 200 milliseconds, in both conditions, the difference in the proportion of fixations to the competitor and distractor was not significant (same-gender pairs: $F_1[1, 38] = 1.281$, $p = 0.272$; $F_2[1, 28] = 0.433$, $p = 0.521$; different-gender pairs: $F_1[1, 38] = 3.339$, $p = 0.083$; $F_2[1, 28] = 1.852$, $p = 0.195$), thus allowing an offhand comparison of the time-frame from 200 to 600 ms.

In the case of the same-gender pairs, we expected to observe a competition effect, just as we did with the native listeners, since the non-native listeners would also be confronted with the fact that the article preceding the noun in the instruction agreed with both the target and competitor nouns, so that there would be no way for them to know which noun they would be asked to click on before the disambiguation point.[10] This, again, was confirmed by the results of the ANOVA from 200 to 600 ms, which showed that the competitors were fixated significantly more often than the distractors ($F_1[1, 38] = 17.630$, $p = 0.000$; $F_2[1, 28] = 11.535$, $p = 0.004$).

In the case of the different-gender pairs also, our predictions were the same as with the native listeners: given that the article in the instruction agreed with the target but not with the competitor, either the participants would use gender information to make an early guess at which object would be the target (in which case we should not observe competition), or they would not, resulting in competition between target and competitor nouns until the disambiguation point was reached. Interestingly, however, in this experiment, we observed that the non-native listeners behaved differently from the native listeners in Experiment 1b, for here the statistical analysis showed a significant difference between the fixations to the competitors and to the distractors ($F_1[1, 38] =$

[10] Moreover, the article preceding the noun in the instruction also agreed with the German translations of the target and its competitor: in *"Cliquez sur le fil... /m/"*, the French masculine article also agrees with the German nouns $Film_{[masc]}$ and $Filter_{[masc]}$.

9.152, $p = 0.007$; $F_2[1, 28] = 7.115$, $p = 0.018$), thus showing that the target and the competitor nouns both initially entered the competition set.[11]

Thus, the results of the experiment seem to show that, even when, in the spoken instruction, the gender-marked article preceding the target noun did not agree with the competitor noun in French, our non-native participants were unable to make use of this fact to resolve competition and thus early on direct their attention towards the target noun.

This is all the more surprising given that it cannot be attributed to a lack of knowledge of the French gender of the competitor nouns. Indeed, after completing the eyetracking experiment, the German-speaking participants were asked to fill in a written questionnaire, in which they were asked to circle the correct article (*le* or *la*) for each target and competitor noun in the experiment. Overall, the participants circled the correct article for 92 % of the nouns (93 % of the target nouns, which they had just heard in the instructions of the experiment, and 91.2 % of the competitors, which were not mentioned during the experiment), thus guaranteeing that, most of the time, our Germanophone listeners knew the correct gender of the competitor in French, although they seemingly did not make use of it during spoken word-recognition.

This observation, then, left two interpretations open: if non-native listeners knew the gender of the nouns, but did not put it to use, were they completely ignoring gender information, or were they using their mother-tongue's gender? Indeed, in our experiment, the French article in the spoken instruction *did* agree in gender with the German translation of the competitor noun (e. g., the article la_{FEM} in *"Cliquez sur la ca. . . [ssette]"* agreed, for instance, with the gender of the German competitor $Kanone_{[fem]}$). Thus it was possible that the Germanophones were using German gender information encoded in their native lexicon, or, alternatively, it was possible that they were using no gender information at all—in which case we would also expect to see competition, given the onset overlap of target and competitor nouns.

[11] Note also that the somewhat longer putative overlap of the onset of the target and competitor nouns in the different-gender pairs (mean for the same-gender pairs: 128 ms; for the different-gender pairs: 150 ms) would, if anything, possibly augment the competition effect in this condition, but that we nonetheless get a larger competition effect with the same-gender pairs.

Experiment 2 was run in order to tease these alternatives apart: given that in our different-gender pairs, target and competitor differed in gender in French, but not in German, by simply translating the spoken instructions into German, we would obtain a set up in which target and competitor nouns did not differ in gender on the surface, although they did differ in the underlying translations. By running this experiment with Francophone native listeners who were proficient in German, we would then be able to observe whether the participants showed competition in both same-gender and different-gender pairs, as would be expected if they ignored gender altogether, or whether they made use of their mother-tongue's gender, thus wrongly reducing competition in the case of the different-gender pairs.

Before presenting the results obtained with Francophones listening to German, however, we will first present those of a control group of German native listeners, which was tested in order to confirm the hypothesis that in German, our materials should elicit competition in both conditions, due to the absence of superficial gender difference between targets and German translations of the competitors in any of the two conditions.

3 Experiment 2: German Instructions

Method

Participants

Experiment 2a: Native Listeners. Sixteen Germanophones participated in Experiment 2a, and were paid 5 euros in return. They were all students (mean age of 22), and had normal or corrected-to-normal vision and normal hearing. Most were born and raised in Germany, but one was the son of a German diplomat and born in Greece, and another was born in Brussels and grew up there until age 13. Apart from the participant who was born in Brussels, only 3 others had made long stays in French-speaking countries (2 years on average). Many had learned French as a foreign language in school, but they were not required to exercise their proficiency here (they were not at all aware of the import of French in the experiment).

Experiment 2b: Non-Native Listeners. Twenty-five Germanophones, students for the most part (mean age: 22), took part in Experiment 2b in exchange for a compensation of 5 euros. All participants had normal hearing and normal vision, except for one who wore very thick glasses. All were raised in France from early childhood (6 months) on, except for 2 from French-speaking parts of Cameroon and one from French-speaking Belgium. However, 4 turned out to be bilinguals, having been taught German as children by one of both parents or by close relatives.[12]

On average, the non-bilingual participants had studied German as a foreign language for 10 years, starting at a mean age of 12. At the time of the experiment, all were living and studying in Germany; 8 had already been there for

a rather long period (3 years on average). 10 explicitly stated that they could speak German fluently and understand it without difficulty.

Materials

The target and competitor nouns in this experiment were the same as in Experiment 1 (see the list in the Appendix on page 79). Note that this means that in the language which was spoken during the experiment, namely German, there was no gender opposition between target and competitor in any of the items (e. g. *Kassette* and *Kanone* are both feminine); it was only in the underlying French translations of the materials that the competitor (e. g. $canon_{[masc]}$) did not match the gender of the target (e. g. $cassette_{[fem]}$) in what we have previously called "different-gender" noun pairs. However, since in the following we are going to investigate the reactions of Francophone learners of German, and we expect that the mother-tongue of the non-native listeners may have an influence on how they recognize words in a foreign language, we will continue here to make a distinction between "different-gender" and "same-gender" pairs (i. e. target-competitor pairs in which, in German as in French, the competitor was of the same gender than the target, such as $Filter_{[masc]}/filtre_{[masc]}$, which was paired with the target $Film_{[masc]}/film_{[masc]}$).

Distractor and filler items were also identical to those of the first experiment, and the same experimental displays were used.

The German target nouns were embedded in German carrier sentences which asked the participants to identify one particular picture among the four in display, such as for example *"Wo befindet sich* die_{FEM} *Kassette?"* (English: "Where is the tape [located]?"). Again, the target nouns were preceded by an agreeing definite article: *die* (/diː/) for feminine nouns, *der* (/deɐ̯/) for masculine ones, and *das* (/das/) for some neuter filler items. An interrogative carrier sentence was chosen so that the nominal phrase containing the target noun would appear at the end of the sentence, and at the same time be in

[12] Although we specifically advertised for participants who were not bilinguals, some did end up participating, probably because they did not consider their German as good as that of other native speakers, although they grew up with one parent or a close relative speaking German. The questionnaire on their linguistic background which we gave them after the eyetracking part of the experiment helped us sort this matter out.

the nominative case, because it was the subject of the sentence; the reason for this was that nominative is usually the first form learned by students of German as a foreign language, and it is also commonly the unmarked ground form in they study their vocabulary. Thus, by comparison with declined forms, this sentence offered the highest chance that the non-native participants in Experiment 2b would be able to very quickly extract gender information from the article preceding the noun.

The spoken instructions were recorded under the same conditions as in Experiment 1 by a Germanophone native speaker and then cut into separate files; a 500 ms blank was added at the beginning, the onset times of the target nouns and the duration of the putative overlap between target and competitor were measured: on average, the overlap was 200 ms for the same-gender pairs and 174 ms for the different-gender pairs.

Procedure

Overall, the procedure was exactly similar to that in Experiment 1. This time, however, the experiment was conducted entirely in German with all participants, be they native speakers or not. The written instructions containing the explanation of the task were translated into German on the basis of those used in Experiment 1. Although the non-native participants *did* know from our advertising that the experiment concerned foreign language processing, the Germanophone natives were in no way made aware of the potential cross-language implications of the experiment.

For the non-native speakers, the linguistic background questionnaire after the experiment was in French, but instead of the French nouns used in Experiment 1, the list in which they were asked to indicate the correct gender of the nouns contained all of the German targets and competitors. In order to make our objective less conspicuous, one third neuter nouns were added to the list, namely the neuter nouns used as targets and distractors in the filler items, so that for each noun, the participants had to circle either *der*, *die* or *das* as the correct nominative definite article for that noun.

Moreover, for the Germanophone listeners, the first part of the questionnaire (i. e. on their linguistic background, what languages they had studied, etc.) was translated into German, and their was no gender quiz at the end (given that

		Targets	Competitors	Av. Distractors
0–200 ms	same-gender pairs	19.2 %	21 %	17.7 %
	different-gender pairs	23 %	18 %	19 %
200–600 ms	same-gender pairs	34.1.5 %	25.7 %	11.7 %
	different-gender pairs	41.4 %	21 %	11.5 %

Table 4: Average proportions of fixations for each type of picture (targets, competitors and averaged distractors) in Experiment 2a, over the time-windows from 0 to 200 ms and from 200 to 600 ms, in both conditions.

we assumed that they knew the correct gender of nouns in their own mother-tongue).

The data was coded in a similar fashion to Experiment 1, and again, averaged proportions of fixations to each type of picture for each condition were computed by slices of 10 milliseconds, in separate analyses by participants and by items.

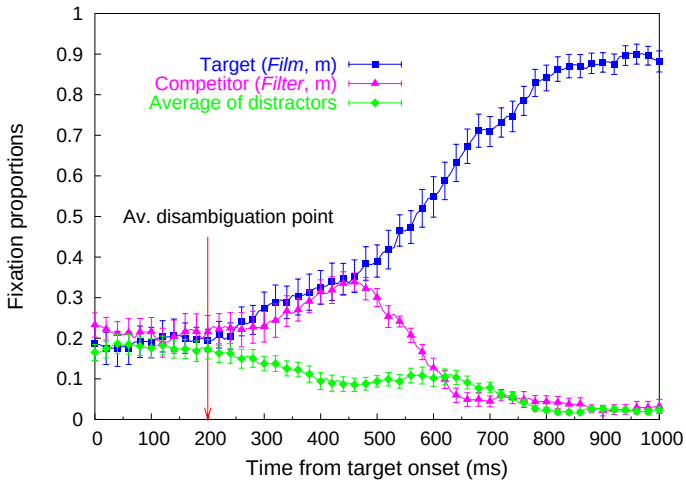
Results

Experiment 2a: Germanophones Listening to German

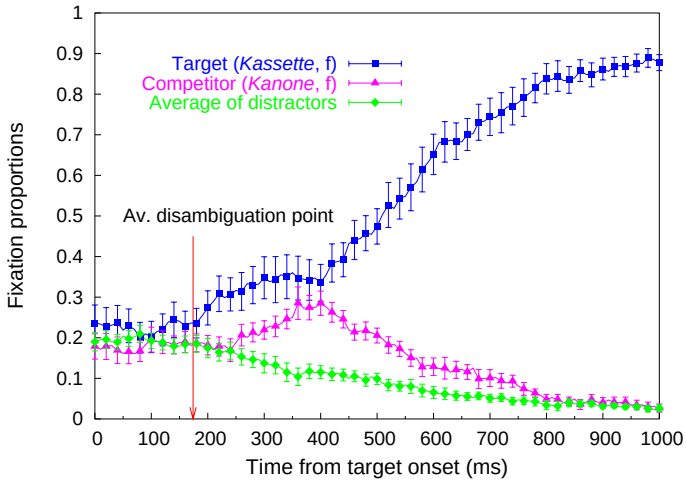
One participant, who used peripheral vision (see on p. 54), as well as the participant born and raised in Brussels, which we considered as a bilingual, were excluded from the analysis. Moreover, the data from 2 trials where we had a technical failure (no sound or full hard drive) was also removed (in the second case, we continued the experiment after remedying the situation).

Figure 6 on page 63 presents the averaged proportions of fixations by participants, from 0 to 1000 ms after target onset, first for the same-gender pairs (Figure 6a) and then for the different-gender pairs (Figure 6b). As previously, the fixation proportions for the two unrelated distractors were averaged. Competition was observed in both conditions. In the case of the same-gender pairs, the gap between the proportions of fixations to the competitors and to the distractors probably becomes significant shortly before 300 ms, point at which the standard error bars do not overlap anymore, and the curves rejoin around 600 ms. In the case of the different-gender pairs, both lines separate even a little earlier, at about 250 ms, and remain apart until after 700 ms.

We conducted the same statistical analyses as in Experiment 1, averaging for each condition, by participants and by items, the fixation proportions to each picture type over the time-windows extending from from 0 ms to 200 ms and



(a) Same-gender pairs



(b) Different-gender pairs

Figure 6: Averaged proportions of fixations by participants in Experiment 2a, from 0 to 1000 ms after target onset. The vertical bars show the standard error.

from 200 ms to 600 ms after target onset. Mean fixation proportions overall are shown for each time-frame in Table 4.

One-factor ANOVAs with picture type (competitor or distractor) as within-factor were computed on the by-participant and by-items averages over each time-frame. Over the first 200 ms after target noun onset, the analyses showed no reliable differences in fixations to the competitor and to the distractor (same-gender pairs: $F_1[1, 26] < 1$, $p = 0.441$; $F_2[1, 28] = 1.192$, $p = 0.293$; different-gender pairs: $F_1[1, 26] < 1$, $p = 0.722$; $F_2[1, 28] < 1$, $p = 0.836$). This means that prior to the point where fixations could be affected by acoustic input from the target, both the competitor and the distractor were fixated equally often, thus excluding a general bias toward the competitor pictures and allowing a direct comparison over the subsequent time-frame.

Recall now that in the language spoken during the experiment, German, the gender of the competitor matched that of the target in both same-gender (e. g. $Film_{[masc]}/Filter_{[masc]}$) and different-gender pairs (e. g. $Kassette_{[fem]}/Kanone_{[fem]}$); the term "different-gender" referred only to the gender of the competitor in the underlying French translations of the nouns. Consequently, since even those participants which by chance happened to know French were not at all aware of its import, our predictions were the same for both conditions: namely, we expected to observe competition in both cases, since German gender information could not be used in any way to exclude the competitor as a potential lexical candidate. This was confirmed by the results of the two ANOVAs which we ran from 200 to 600 ms: the difference between the proportions of fixations to the competitor and to the distractor was significant for both same-gender ($F_1[1, 26] = 22.378$, $p = 0.000$; $F_2[1, 28] = 15.257$, $p = 0.002$) and different-gender pairs ($F_1[1, 26] = 14.319$, $p = 0.002$; $F_2[1, 28] = 11.052$, $p = 0.005$).

Thus, as we anticipated, it appeared that, in the German version of the experiment, the gender information available from the article did not support a reduction of the competition set in neither same-gender nor different-gender pairs, due to the absence of gender difference in the German competitors.

Moreover, note in passing that the behavior of the Germanophone natives in this experiment was also exactly the same as that of the Germanophone learners of French in Experiment 1b, thus asking once again the question of whether non-native listeners use the gender of their mother-tongue in foreign

language processing (since our Germanophone listeners apparently reacted in the same way in native and non-native processing), or whether they use no gender information at all (in which case the observations in Experiment 1b and Experiment 2a would be similar, although they would arise from different sources). This matter will now finally be approached in Experiment 2b.

Experiment 2b: Francophones Listening to German

The 4 bilingual participants were removed from the analysis, because their behavior might be different from that of non-native listeners: the single analysis of the gender questionnaire revealed that they made distinctly less mistakes than the non-natives, thus suggesting that they might be capable of a more native-like type of processing. Moreover, the data from the one participant who wore very thick glasses had to be removed because it appeared to be unreliable (many fixations to the grey background and hardly any to the pictures), possibly due to technical difficulties with the calibration of the eyetracker.

Two trials in which no sound was played, as well as a few trials in which the Francophone speakers had not clicked anywhere or clicked on the wrong picture (3 %) were also eliminated, thus leaving us with fixation data from 291 same-gender trials and 291 different-gender trials. The low percentage of errors suggests that the participants had no difficulty performing the task in German.

Figure 7 on page 67 presents the averaged proportions of fixations by participants, from 0 to 1000 ms after target onset, first for the same-gender pairs in Figure 7a and then for the different-gender pairs in Figure 7b. Fixation proportions for the two unrelated distractors were again averaged. This time, competition was only observed in the same-gender pairs (Figure 7a), where the proportion of fixations to competitors and distractors diverge shortly before 200 ms after target noun onset, and the curves meet again somewhere between 550 and 650 ms. On the contrary, in the different-gender pairs (Figure 7b), the two lines remain together until about 500 ms, which is long after the point where the proportion of fixations to the target start rising.

Average proportions of fixations to competitors and distractors over the time-windows from 0 to 200 ms and from 200 to 600 ms are shown in Table 5. As before, we conducted one-factor ANOVAs on the mean proportions of fixations over each time-frame, with picture type as within-factor, and again, fixations to the competitors and distractors did not differ in either of the initial regions going

		Targets	Competitors	Av. Distractors
0–200 ms	same-gender pairs	29.4 %	20.7 %	17.4 %
	different-gender pairs	22.4 %	20 %	19.5 %
200–600 ms	same-gender pairs	32.5 %	23.3 %	15.6 %
	different-gender pairs	33.1 %	19.6 %	15.8 %

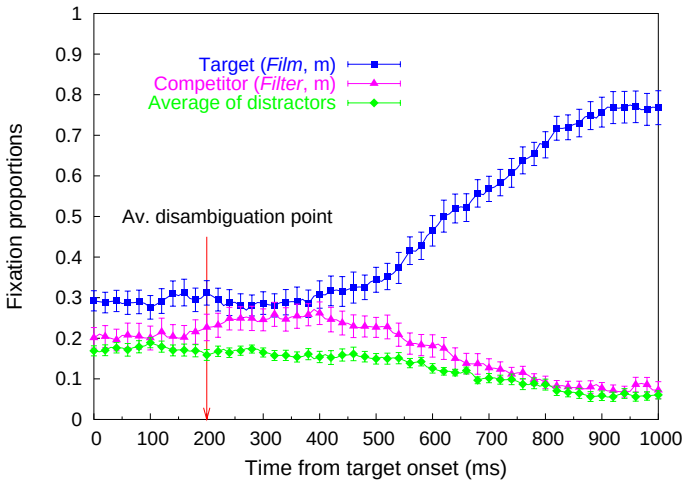
Table 5: Average proportions of fixations for each type of picture (targets, competitors and averaged distractors) in Experiment 2b, over the time-windows from 0 to 200 ms and from 200 to 600 ms, in both conditions.

from 0 to 200 ms (same-gender pairs: $F_1[1, 38] < 1$, $p = 0.337$; $F_2[1, 28] < 1$, $p = 0.455$; different-gender pairs: $F_1[1, 38] < 1$, $p = 0.895$; $F_2[1, 28] < 1$, $p = 0.865$), so that we could casually compare the further course of the fixation proportions.

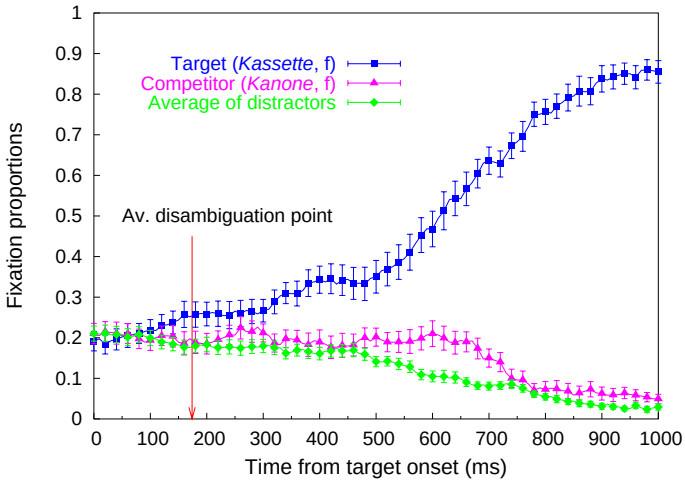
By comparison with Experiment 2a, our predictions this time were different: although if we had considered only the language spoken during the experiment, the predictions would have remained the same as before, this time we actually considered it possible that the mother-tongue of the Francophone listeners might have an influence on their processing of German. Thus, in this experiment, our expectations for the same-gender and different-gender pairs were different, and, crucially, they were based on characteristics of the participants' mother-tongue, which was *not* spoken during the experiment.

Concerning the same-gender pairs, we hypothesized that, since target and competitor nouns shared gender in both languages (e. g., $Film_{[masc]}$ was paired with the German $Filter_{[masc]}$, whose French translation, *filtre*, was also masculine), neither German nor French gender information would constrain initial competitor activation, thus revealing a competition effect, which indeed was found: $F_1[1, 38] = 8.438$, $p = 0.009$; $F_2[1, 28] = 10.206$, $p = 0.006$.

In the different-gender pairs, however, target and competitor also shared gender in German (e. g. $Kassette_{[fem]}/Kanone_{[fem]}$), but they were of different gender in the underlying French translations (competitor: $canon_{[masc]}$). Thus whereas German gender could not exclude the competitor early on from lexical competition, French gender information might, because the article in the spoken instruction would not agree with the French competitor: for example, in *"Wo befindet sich die*$_{\text{FEM}}$ *Ka. . . "*, the feminine article agrees only with German *Kassette* and/or French *cassette*, but not with *canon*. If the Francophone listeners ignored gender information, we should observe competition, while if



(a) Same-gender pairs



(b) Different-gender pairs

Figure 7: Averaged proportions of fixations by participants in Experiment 2b, from 0 to 1000 ms after target onset. The vertical bars show the standard error.

they made use of gender information from their mother-tongue, we should not (note that we already knew from Experiment 1b that it was unlikely that the non-native listeners would be able to use German gender information, but if this were the case, it would be impossible to tell it apart from a situation in which the non-natives used no gender information at all).

It turned out that the ANOVA was non-significant ($F_1[1, 38] = 2.429$, $p = 0.136$; $F_2[1, 28] = 2.201$, $p = 0.160$): despite the onset overlap of target and competitor, the competitor was not activated when the article in the spoken instruction did not agree in gender with the French translation of the competitor. Apparently, although they were listening to a foreign language, the non-natives drew on gender information in their mother-tongue, thus reducing competition when they should not have, because in German, the gender of the article should not have excluded the competitor as potential lexical candidate.

These results thus clarify those of Experiment 1b: although earlier we could not exclude the possibility that non-native listeners used no gender information at all in foreign language processing, we now know that they actually use their mother-tongue's gender information. This, then represents a case of morphosyntactic "capturing", since nouns whose gender is different in the mother-tongue and the foreign-language (for example $Kanone_{[fem]}$, which in French is $canon_{[masc]}$) are "captured" by the native gender category (here masculine), and treated as if they were part of the wrong gender category.

Again, as in Experiment 1b, we additionally verified that the fact that the non-native participants used their mother-tongue's gender could not be attributed to their not knowing the correct gender for the German nouns: after the eyetracking, they underwent a vocabulary test in German, in which they had to name the correct gender for each target and competitor noun in the experiment. Overall, the average score was 78.8 % correct (81.3 % for the targets, which the participants had just heard during the experiment, and 76.3% for the competitors, which they had not heard), although there were only 7 % confusions between masculine and feminine (which were the only ones relevant to the experiment),[13] thus guaranteeing that the morphosyntactic "capturing" effect could not be ascribed to a lack of knowledge of the language under investigation on the part of the participants.

[13] In the gender test, the Francophone participants confused German masculine and neuter much more often (10.3 % of the nouns) than feminine and neuter (3.75 %) or masculine and feminine (7 %).

4 General Discussion

The experiments presented in this thesis reveal that native gender categories can interfere with spoken word-recognition in foreign languages—at least when the foreign language also has gender and the categories in both languages share many nouns but not all. Moreover, the results of the experiments also allow us to pinpoint the origin of the gender effect described in many offline studies (see Section 1.3) and in Dahan et al. (2000).

Using eyetracking in visual worlds, we presented participants with displays composed of four pictures on a computer screen and asked them to click on one of the pictures. In each experimental display, two of the pictures, the target and the competitor, overlapped in onset. We expected that, other factors notwithstanding, the competitor would initially be activated alongside the target due this similarity. By interpreting eye-movements as evidence for the activation of the words corresponding to the pictures, we then compared the activation of the competitor with the activation of words unrelated to the target (termed 'distractors'), in order to determine whether or not the competitor was activated in various conditions.

There were two sets of instructions: in French for Experiment 1 and in German for Experiment 2. In both cases, the critical noun in the instructions was always preceded by an agreeing, singular, gender-marked article. However, the items differed in that depending on the language of the experiment and on the experimental condition, in some cases, it was possible that the gender information carried by the article would be used to exclude some nouns from the competition set, whereas in other cases, the article agreed with all the relevant nouns and thus could not be used in this manner.

In Experiment 1, we compared the reactions of Francophones and proficient Germanophone learners of French, using the same instruction as Dahan et al. (2000): *"Cliquez sur* $le_{\text{MASC}}/la_{\text{FEM}}\ldots$*"*. Half of the items were based on target-competitor pairs of the same gender ("same-gender" pairs), whereas the other

half (the "different-gender" pairs) were based on pairs in which the gender of the competitor differed from the gender of the target in French, but not in German. In the case of the native listeners, the results showed competition in the case of the "same-gender" pairs, but none in the case of the "different-gender" pairs: the Francophones listeners appeared to be able to use gender to suppress competition when appropriate. However, in the case of the non-natives, we observed competition in both conditions, thus revealing that the Germanophone participants were unable to make use of gender in the same way as the natives. Two interpretations of this observation were then possible: were the non-native participants completely ignoring gender, or were they using the gender categories of their mother-tongue?

This question was subsequently addressed in a German version of the same experiment, Experiment 2, using the same materials. This time, we tested Germanophone natives and Francophone learners of German. The instruction-sentence was *"Wo befindet sich der*$_{\text{MASC}}$*/die*$_{\text{FEM}}$*..."* (English: "Where is the..."). On the surface, all noun pairs had the same gender; it was only in the underlying French translations of the competitors that there was a difference between "same-gender" and "different-gender" pairs. As a consequence, the Germanophone participants showed the expected competition effects in both conditions, but interestingly, the Francophone participants did not show competition in the "different-gender" pairs, thus suggesting that they were making use of French gender information even while listening to German.

4.1 Replicating Dahan et al. (2000)

In Section 1.5, we briefly described the Dahan et al. (2000) study, which, building on the linking hypothesis proposed by Allopenna et al. (1998), was the first to apply the high time-resolution available with eyetracking to the further examination of the facilitation effect found with gender in many other studies of spoken word-recognition.

As a baseline, the first experiment in this thesis replicated the results obtained by Dahan et al. (2000), with a slightly different setup: whereas Dahan et al. presented the same materials in the singular and in the gender-unmarked plural to two different groups of participants, we used the singular in all cases, and

contrasted instead in a single experiment target-competitor pairs with matching gender and target-competitor pairs with differing gender.

Our results with French natives in Experiment 1a exhibited competition in the case of the same-gender pairs but no competition in the case of the different-gender pairs, thus confirming Dahan et al. (2000)'s conclusion that it is possible for the initial set of lexical candidates to be constrained by the preceding context: in the absence of preceding gender information, competition between the target and competitor nouns took place in French just as in English (see, for example, Tanenhaus et al., 1995), but when gender could help to determine the target before the point where the target noun itself was unambiguous, then gender information preceding the noun was used by native listeners to reduce competition during spoken word-recognition.

Taking up the example from the beginning of the introduction (see p. 17), this would mean that a Francophone native listener could from the very start reduce the set of lexical candidates by about half: instead of 2326 French nouns beginning with /k/, there would only 917 to pick from after a feminine article, instead 623 with the onset /ka/, only 218, and instead of 34 starting with /kas/, only 18, thus resulting in faster and easier recognition of the target word *"la cassette"*, since it has been shown that spoken word-recognition is slower when there is more competition among the members of the initial set of lexical candidates (Norris et al., 1995; Vroomen & de Gelder, 1995).

These findings seem to correlate very well with Grosjean et al. (1994)'s observations concerning the completion candidates proposed by participants in their gating task: when a French nominal phrase was preceded by an agreeing article, the number of different proposals was much lower, and the target word occurred at an earlier gate than when the article was absent.

Moreover, they also constitute an explanation of the facilitation and inhibition effects found by the various offline studies we described in Section 1.3. Indeed, assuming that gender information can be used to reduce the initial set of lexical candidates means that when gender information is available, recognition should proceed more rapidly and require less effort. Moreover, on the opposite, when incorrect gender information is carried by the preceding context, it would then lead to the activation of an incorrect set of lexical candidates, in which the target word would not be contained, with the consequence that

spoken word-recognition would proceed more slowly, because the target word would have to be re-activated at some later point in time.

4.2 A Morphosyntactic "Capturing" Effect

The pattern observed with non-native listeners of French (Experiment 1b), however, was altogether different: although they were proficient in French and knew the gender of the nouns in the experiment (given a list of the experimental nouns, they were able to circle the correct gender for 92 % of the nouns; see p. 57), our Germanophone learners of French did not appear to be able to use gender in the same way as the French natives, thus experiencing competition with both same-gender and different-gender pairs.

Additionally, the results of Experiment 2b demonstrated that Francophone learners of German, too, did not behave like Germanophones: although Germanophones showed competition in both conditions, native speakers of French listening to German appeared to suppress competition in the case of different-gender pairs. As in Experiment 1b, this could not be attributed to a lack of knowledge of the gender of German nouns, for the participants rarely confused masculine and feminine nouns in the post-hoc offline gender test (see p. 68).

Instead, our Francophones who had learned German happened to show just the same pattern as when Francophones were listening to their own mother-tongue in Experiment 1a. This leads us to believe that they were actually using their native gender categories while processing German, by mapping French gender categories onto German input. Moreover, it is rather interesting to note that the non-native participants in Experiment 1b also exhibited an identical pattern to that of the native Germanophone participants in Experiment 2a, which might also be explained by venturing that they were making exclusive use of German gender categories.

Thus, it seems that on the morphosyntactic level, as on the phonetic level (Weber & Cutler, 2004), native categories have a tendency to "capture" non-native input: when hearing the beginning of *"la ca[ssette]"*, for example, it seems that Germanophone learners of French are able to extract the information that the article is feminine astonishingly fast for non-native listeners, but then they appear to activate all feminine nouns matching the noun onset in their mother-tongue (together with any matching French nouns), including nouns whose

French translation may not match in gender, such as $Kanone_{[fem]}$ (French: $canon_{[masc]}$).

Conversely, when hearing the beginning of *"die Ka[ssette]"* in German, Francophone learners also notice the article is feminine. Were it not for gender, they would then probably activate all nouns beginning in /ka/ in both French and German (given the results of Weber & Cutler, 2004), but instead they make use of gender information in their mother-tongue and increase the activation of native feminine words (possibly also inhibiting masculine words), so that due to the relation between translation pairs such as *canon*/*Kanone*, the effect of the mother-tongue's gender ends up impinging on spoken word-recognition in the foreign language.

Why, however, do the non-natives listening to German not make use of German gender, and thus activate *Kanone*, since the gender test after the eyetracking experiment proved that they really *do* know the gender of the German nouns used? Indeed, remember that in Section 1.1, we assumed that the process of spoken word-recognition is similar in native and non-native languages, particularly in the light of the competition effects in non-native processing evidenced by Spivey & Marian (1999).

Here, it should be noted that it is not because the basic mechanism underlying spoken word-recognition, namely competition between lexical candidates, is similar in native and non-native processing that all factors which play a role in spoken word-recognition must also act in the same way when processing our mother-tongue or a foreign language: it is very much possible that among the many factors known to influence spoken word-recognition, some act in a similar fashion and some do not. Among these factors, the capacity to use the preceding context to constrain the initial set of lexical candidates, demonstrated in native processing by Dahan using gender, must not necessarily apply also to non-native processing: it could be that non-native listeners, although they know the gender of a noun, cannot make use of it in spoken word-recognition, or that gender knowledge is available too slowly, so that non-natives were able to access it only in the offline questionnaire, when they had time to reflect and the entire word (including its ending) was available.

In any case, it seems that due to the confusion created by the partial overlap of French and German gender categories, non-natives appear to stretch the gender categories of the foreign tongue so that they contain gender-matching nouns

in their mother-tongue, so that native gender categories end up overruling the gender categories of the non-native language, with the unfortunate consequence of sometimes possibly leading to an incorrect set of initial lexical candidates, and thus delayed and more effortful spoken word-recognition.

4.3 The Origin of the Gender Effect: Gender Categories

The results we obtained with non-native listeners concerning potential gender effects on spoken word-recognition are all the more interesting because they now allow us to venture a guess at whether the origin of the gender effect is form-based or grammar-based.

Indeed, remember that in the introduction we detailed what patterns of results would be expected when listening to a foreign language depending on whether the gender effect were due to the frequent occurrence of given phoneme sequences before particular nouns, or to the frequent occurrence of gender-matching determiners before nouns.

We also mentioned an unpublished study by Dahan and colleagues in which the preceding gender-carrying context (a gender-marked pre-nominal adjective) had a lower frequency of co-occurrence with the noun. Dahan and colleagues found no reduction of competition due to gender, thus suggesting that the gender effect in the second experiment of Dahan et al. (2000) might have be superficial. On the other hand, Jakubowicz & Faussart (1998) *did* find a facilitation effect with lower frequency French determiners (demonstratives and possessives, in addition to articles) and various interposed adjectives before the noun, thus pointing to a grammar-based explanation.

As in Jakubowicz & Faussart (1998), our results indicate that the gender effect rather originates on the grammatical level of language processing. If the gender effect had been form-based, we would have expected the non-native participants in our experiments to either behave like the native participants, if they rapidly learned the gender categories of the non-native tongue, or to exhibit competition in all cases because gender had no influence whatsoever (the competition set being then determined only by the onset of words in the lexicon). None of these patterns was however observed. Alternatively, if the gender effect were grammar-based, we could also have expected the two same

behaviors, to which a third might be added: in Experiment 2b, it was also possible that the Francophone learners of German would use their native categories to suppress competition in the different-gender pairs, although they should not. As it turned out, this was what we observed, thus allowing us to conclude that the gender effect described in the many studies we reviewed in the introduction must be grammar-based, i. e. due to the mother-tongue's gender categories.[14]

4.4 Conclusion and Future Work

In this thesis, we took a look at the impact of lexical gender on non-native spoken word-recognition. Using eyetracking in visual worlds, we conducted two experiments, in which we compared proficient learners of French and German with respective groups of native listeners.

It appeared that, although gender information on a preceding article can be used by native listeners to constrain the set of potential lexical candidates for recognition, resulting in faster spoken word-recognition, non-native listeners could not use gender information in the same way, even when they theoretically knew a noun's gender. More specifically, it seems that non-native listeners actually make use of the gender categories in their mother-tongue when processing a foreign language, especially when the categories of both languages overlap.

Additionally, our results allowed us to assert that the gender effect, which has been observed in many studies of native spoken word-recognition, is grammar-based, because otherwise there would be no explanation why non-native listeners use the gender categories of their mother-tongue in foreign language processing. This would seem to favor a non-modular account of spoken word recognition, for if spreading activation within the lexicon is excluded (as argued

[14] This is also the conclusion reached by Sekerina (2003), who recently conducted another study of gender effects with eyetracking in visual worlds, using gender-marked color adjectives in Russian. Whether the target noun immediately followed the gender-marked adjective, or whether they were separated by an intervening word (namely a verb, since Russian allows scrambling), the results both showed that gender could be used to facilitate the identification of the referent, the gender-matching distractor being fixated more often than the gender-mismatching distractor: given the target "red_{FEM} $\text{car}_{[fem]}$", the picture of a "red $\text{squirrel}_{[fem]}$" received more fixations than that of a "red $\text{flower}_{[masc]}$". However, it seems that the displays in this experiment did not contain a lexical competitor overlapping in onset with the target, but instead that all the nouns were unrelated to the target (i. e. they were what we have called 'distractors' in this thesis), so that the results may have more to do with the interaction of color and gender in their influence on reference resolution than with lexical access in spoken word-recognition.

by Dahan et al., 2000), then this means that grammatical information about the gender of a word must come from the grammatical level of language processing.

One limitation of the experiments we have presented comes from the fact that, in testing non-native participants, we did not distinguish between levels of proficiency. Thus, it is possible that, although participants who had a very good school-knowledge of their foreign language retained the use of their mother-tongue gender categories in non-native processing, participants who had spent a long time in the foreign country may have gradually learned to use the categories of their foreign language, but that this may have been drowned in the statistical analysis.

Given our observations, it would now also be interesting to see what would happen when non-native listeners are faced with other kinds of interactions between the gender of their mother-tongue and that of their foreign language. We do know already that English natives learning French make no use at all of French gender categories in spoken word-recognition (Guillelmon & Grosjean, 2001), even when they know the gender of French nouns. But what about non-native listeners whose mother-tongue has gender, but where the categories are wider apart than masculine and feminine in French and German? For example, what happens when a foreign category does not exist in the mother-tongue, such as neuter in French? Would French natives be able to learn to use neuter, because it did not interfere with any native category, would they not use it all, or would they confuse it with French masculine, as prognosticated by Corbett (1991)'s affirmation that masculine and neuter have merged in French over the course of history, as well as by our French native participants' high number of confusions between German masculine and neuter in the offline test?

One might also want to research how bilinguals react to gender in spoken word-recognition. Indeed, in our experiments we removed any data from bilinguals from the analysis, because they might have behaved differently from non-native listeners. In particular, in Experiment 2b, the data from the 4 bilinguals seemed to show lexical competition in those items where late learners of German had wrongly suppressed it. Although at first sight this seems to be the same pattern as the native listeners, this may not generalize: it may be the case that bilinguals, on the whole, behave like monolinguals, or it may not.

In native processing, another point would be to look at the influence of other kinds of morphosyntactic information than gender, and at the influence of gen-

der in other languages. In languages with very different gender systems, such as those based on semantics for example, the influence of gender on spoken word-recognition might be different. Additionally, given that our results have confirmed Dahan et al. (2000)'s demonstration that the set of lexical candidates initially considered for spoken word-recognition can be constrained by gender, it would be worthwhile to think about what other morphosyntactic and/or syntactic information (e. g. word category) might also play a similar role, and to approach such questions with online, high-resolution investigation methods such as eyetracking.

Appendix: Experimental Stimuli

Table 1 on page 80 shows the experimental stimuli (targets, competitors and distractors) used in the "different-gender" condition, and Table 2 on page 81 those used in the "same-gender" condition. The targets and competitors are accompanied by their phonetic transcriptions, which follow Le Robert (1995) and Duden (1990), with slight adaptations. The letters "F", "G" and "E" on the left-hand side stand for French, German, and English (translations). In the "different-gender" condition, the French competitor, whose gender does not match that of the target, is printed in bold-face.

Table 1: Different-Gender Pairs

		Targets		Competitors			Distractors	
Feminine Targets	F	carafe	/karaf/	**carrelage**	m	/karlaʒ/	nounours	poisson
	G	Karaffe	/karafə/	Kachel	f	/kaxl̩/	Teddybär	Fisch
	E	“decanter”		“tile”			“teddy bear”	“fish”
	F	carotte	/karɔt/	**capuchon**	m	/kapyʃõ/	cerf-volant	avion
	G	Karotte	/karɔtə/	Kapuze	f	/kaputsə/	Drachen	Flugzeug
	E	“carrot”		“hood”			“kite”	“plane”
	F	cassette	/kasɛt/	**canon**	m	/kanõ/	robinet	fauteuil
	G	Kassette	/kasɛtə/	Kanone	f	/kanonə/	Wasserhahn	Sessel
	E	“tape”		“canon”			“faucet”	“armchair”
	F	flûte	/flyt/	**flocon**	m	/flɔkõ/	lapin	arbre
	G	Flöte	/fløːtə/	Flocke	f	/flɔkə/	Kaninchen	Baum
	E	“recorder”		“flake”			“rabbit”	“tree”
	F	gazelle	/gazɛl/	**garage**	m	/garaʒ/	valise	pied
	G	Gazelle	/gatsɛlə/	Garage	f	/garaʒə/	Koffer	Fuß
	E	“gazelle”		“garage”			“suitcase”	“foot”
	F	marionnette	/marjɔnet/	**matelas**	m	/matla/	peigne	poire
	G	Marionette	/marjɔnɛtə/	Matratze	f	/matratsə/	Kamm	Birne
	E	“puppet”		“mattress”			“comb”	“pear”
	F	marmelade	/marmələd/	**masque**	m	/mask/	phare	clou
	G	Marmelade	/marmələdə/	Maske	f	/maskə/	Leuchtturm	Nagel
	E	“marmelade”		“mask”			“lighthouse”	“nail”
	F	papaye	/papaj/	**palmier**	m	/palmje/	vélo	gâteau
	G	Papaya	/papaja/	Palme	f	/palmə/	Fahrrad	Kuchen
	E	“papaya”		“palmtree”			“bicycle”	“cake”
	F	pizza	/pidza/	**pistolet**	m	/pistɔlɛ/	hérisson	camion
	G	Pizza	/pɪtsa/	Pistole	f	/pɪstoːlə/	Igel	Laster
	E	“pizza”		“pistol”			“hedgehog”	“truck”
	F	râpe	/rap/[a]	**rat**	m	/ra/	champignon	bouton
	G	Reibe	/raibə/	Ratte	f	/ratə/	Pilz	Knopf
	E	“grater”		“rat”			“mushroom”	“button”
	F	tasse	/tas/	**tableau**	m	/tablo/	cloche	bague
	G	Tasse	/tasə/	Tafel	f	/taːfl̩/	Glocke	Ring
	E	“cup”		“blackboard”			“bell”	“ring”
	F	vanille	/vanij/	**vase**	m	/vaz/	couteau	chaussette
	G	Vanille	/vanilə/	Vase	f	/vaːzə/	Messer	Socke
	E	“vanilla”		“vase”			“knife”	“sock”
Masc. Targets	F	balcon	/balkõ/	**balle**	f	/bal/	sapin	fromage
	G	BalkoN	/balkɔŋ/	Ball	m	/bal/	Tanne	Käse
	E	“balcony”		“ball”			“pine tree”	“cheese”
	F	ballon	/balõ/	**barbe**	f	/barb/	télévision	trésor
	G	Ballon	/balɔŋ/	Bart	m	/baːɐ̯t/	Fernseher	Schatz
	E	“balloon”		“beard”			“TV set”	“treasure”
	F	coffre	/kɔfr/	**corbeille**	f	/kɔrbɛj/	patin	tabouret
	G	Kofferraum	/kɔfɐraum/	Korb	m	/kɔrp/	Schlittschuh	Hocker
	E	“car trunk”		“basket”			“iceskate”	“stool”

[a]Here we use the front /a/ in the transcription, because the back /ɑ/ is gradually disappearing in French pronunciation, as mentioned by Le Robert (1995, pp. XXI–XXII).

Table 2: Same-Gender Pairs

		Targets		Competitors		Distractors	
Feminine Targets	F	batterie	/batri/	banane	/banan/	cintre	oiseau
	G	Batterie	/batəriː/	Banane	/banaːnə/	Kleiderbügel	Vogel
	E	"battery"		"banana"		"coat hanger"	"bird"
	F	cafetière	/kaftjɛr/	caméra	/kamera/	lit	âne
	G	Kaffeekanne	/kafekanə/	Kamera	/kamɛra/	Bett	Esel
	E	"coffeepot"		"camera"		"bed"	"donkey"
	F	crêche	/krɛʃ/	crème	/krɛm/	chou-fleur	singe
	G	Krippe	/krɪpə/	Creme	/kremə/	Blumenkohl	Afen
	E	"nativity scene"		"cream"		"cauliflower"	"monkey"
	F	crevette	/krəvɛt/	cravate	/kravat/	réveil	bulldozer
	G	Krabbe	/krabə/	Krawatte	/kravatə/	Wecker	Planierraupe
	E	"shrimp"		"tie"		"alarm clock"	"bulldozer"
	F	limonade	/limɔnad/	libellule	/libɛlyl/	cendrier	porte
	G	Limonade	/lɪmɔnaːdə/	Libelle	/libɛlə/	Aschenbecher	Tür
	E	"lemonade"		"dragonfly"		"ashtray"	"door"
	F	perle	/pɛrl/	perruque	/pɛryk/	escargot	ceinture
	G	Perle	/pɛrlə/	Perücke	/pɛrykə/	Schnecke	Gürtel
	E	"pearl"		"wig"		"snail"	"belt"
	F	biceps	/biseps/	bikini	/bikini/	arbre de Noël	pomme
	G	Bizeps	/biːtsɛps/	Bikini	/bikiːni/	Weihnachtsbaum	Apfel
	E	"biceps"		"bikini"		"Christmas tree"	"apple"
Masculine Targets	F	café	/kafe/	cactus	/kaktys/	pince à linge	voiture
	G	Kaffee	/kafe/	Kaktus	/kaktʊs/	Wäscheklammer	Auto
	E	"coffee"		"cactus"		"clothes peg"	"car"
	F	crabe	/krab/	crocus	/krɔkys/	sac à main	bouteille
	G	Krebs	/kreps/	Krokus	/krokʊs/	Handtasche	Flasche
	E	"crab"		"crocus"		"handbag"	"bottle"
	F	dinosaure	/dinɔzor/	diamant	/diamɑ̃/	canne à pêche	pain
	G	Dinosaurier	/dinɔzaurjɐ/	Diamant	/diamant/	Angel	Brot
	E	"dinosaur"		"diamond"		"fishing pole"	"bread"
	F	domino	/dɔmino/	dollar	/dɔlar/	bateau	gland
	G	Domino	/doːmino/[b]	Dollar	/dɔlar/	Schiff	Eichel
	E	"domino"		"dollar"		"boat"	"acorn"
	F	film	/film/	filtre	/filtr/	enveloppe	lampe de poche
	G	Film	/fɪlm/	Filter	/fɪltɐ/	Briefumschlag	Taschenlampe
	E	"film"		"filter"		"envelope"	"torch"
	F	kiwi	/kiwi/	kiosque	/kiɔsk/	nuage	marteau
	G	Kiwi	/kiːviː/	Kiosk	/kiɔsk/	Wolke	Hammer
	E	"kiwi"		"kiosk"		"cloud"	"hammer"
	F	parc	/park/	passeport	/paspor/	église	oeuf
	G	Park	/park/	Paß	/pas/	Kirche	Ei
	E	"park"		"passport"		"church"	"egg"
	F	trône	/tron/	tracteur	/traktɜr/	fourchette	livre
	G	Thron	/troːn/	Traktor	/traktoːɐ̯/	Gabel	Buch
	E	"throne"		"tracteur"		"fork"	"book"

[b] Note that although in the IPA transcription, the first vowel in the French and German words are different, the sounds are in reality not that far away from each other, because the open [ɔ̯] in French is relatively similar in this case to the German closed [o̯] (alternative pronunciation also: /ˈdɔmɪnoʊ/).

Bibliography

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, *38*, 419–439.

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The* CELEX *lexical database* [CD-ROM]. Philadelphia: Linguistic Data Consortium, University of Pennsylvania.

Bates, E., Devescovi, A., Hernandez, A., & Pizzamiglio, L. (1996). Gender priming in Italian. *Perception & Psychophysics*, *85*(7), 992–1004.

Van Berkum, J. J. A. (1996). *The psycholinguistics of grammatical gender.* Nijmegen: Nijmegen University Press. (Doctoral dissertation, Katholieke Universiteit Nijmegen)

Bickford, J. A., & Tuggy, D. (2002). *Electronic glossary of linguistic terms.* (`http://www.sil.org/mexico/ling/glosario/E005ai-Glossary.htm`)

Bidot, E. (1925). *La Clef du genre des substantifs français (Méthode dispensant d'avoir recours au dictionnaire)* [The Key to gender of French common nouns (Method without recourse to the dictionary)]. Poitiers: Imprimerie Nouvelle. (Cited on p. 57 by Corbett, 1991)

Blair, I. V., Urland, G. R., & Ma, J. E. (2002). Using Iternet search engines to estimate word frequency. *Behavior Research Methods, Instruments, & Computers*, *34*(2), 286–290.

Boersma, P., & Weenink, D. (2003). *Praat: doing phonetics by computer* [Computer software]. Freely available from `http://www.praat.org/`.

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*(3), 407–419.

Colé, P., & Segui, J. (1994). Grammatical incongruency and vocabulary types. *Memory and Cognition*, *22*, 387–394. (Cited by van Berkum, 1996, and Guillelmon & Grosjean, 2001)

Content, A., Mousty, P., & Radeau, M. (1990). Brulex. Une base de données lexicales informatisée pour le français écrit et parlé [Brulex. A computerized database with lexical information for written and spoken French]. *L'Année Psychologique*, *90*, 551–566.

Corbett, G. (1991). *Gender*. Cambridge: Cambridge University Press.

Cuetos, F., Mitchell, D. C., & Corley, M. M. B. (1996). Parsing in different languages. In M. Carreiras, J. García-Albea, & N. Sabastían-Gallés (Eds.), *Language processing in Spanish* (pp. 145–187). Mahwah, NJ: Erlbaum.

Cutler, A. (1995). Spoken word recognition and production. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition* (Vol. 11, pp. 97–136). San Diego: Academic Press.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001a). Time course of frequency effects in spoken-word recognition: Evidence from eye-movements. *Cognitive Psychology*, *42*, 317–367. (Cited by Tanenhaus, Magnuson, Dahan, & Chambers, 2000)

Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001b). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, *16*(5/6), 507–534. (Cited by Tanenhaus et al., 2000)

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.

Duden. (1990). *Das Duden Aussprachewörterbuch: Wörterbuch der deutschen Standardaussprache* (third ed.). Mannheim-Wien-Zürich: Duden Verlag.

Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, *7*, 195–225.

Frauenfelder, U. H., & Tyler, L. K. (Eds.). (1987). *Spoken word recognition*. Cambridge, MA: MIT Press/Bradford Books.

Friederici, A. D., & Jacobsen, T. (1999). Processing grammatical gender during language comprehension. *Journal of Psycholinguistic Research*, *28*(5). (Special Issue: Processing Grammatical Gender)

Grosjean, F., Dommergues, J.-Y., Cornu, E., Guillelmon, D., & Besson, C. (1994). The gender-marking effect in spoken word recognition. *Perception and Psychophysics*, *56*(5), 590–598.

Guillelmon, D., & Grosjean, F. (2001). The gender marking effect in spoken word recognition: The case of bilinguals. *Memory and Cognition*, *29*(3), 503–511.

Gurjanov, M., Lukatela, G., Lukatela, K., Savic, M., & Turvey, M. (1985). Grammatical priming of inflected nouns by the gender of possessive adjectives. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(4), 692–701. (Cited by van Berkum, 1996, and Guillelmon & Grosjean, 2001)

Hockett, C. F. (1958). *A course in modern linguistics*. New York: Macmillan. (Cited on p. 1 by Corbett, 1991)

IMSI MasterClips. (1990). *Premium image collection 303,000* [Image database]. (`http://www.imsisoft.com`)

Jakubowicz, C., & Faussart, C. (1998, November). Gender agreement in the processing of spoken French. *Journal of Psycholinguistic Research*, *27*(6), 597–617.

Keller, F., Lapata, M., & Ourioupina, O. (2002). Using the web to overcome data sparseness. In J. Hajič & Y. Matsumoto (Eds.), *Proceedings of the conference on empirical methods in natural language processing* (pp. 230–237). Philadelphia.

Kempen, G. (1996). Computational models of syntactic processing in language comprehension. In A. F. J. Dijkstra & K. J. M. J. de Smedt (Eds.), *Computational psycholinguistics: AI and connectionist models of human language processing.* London: Taylor & Francis.

Marian, V., & Spivey, M. (1999). Activation of Russian and English cohorts during bilingual spoken word recognition. In *Proceedings of the twenty-first annual conference of the cognitive science society* (pp. 349–354). Mahwah, NJ: Lawrence Erlbaum.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 71–102). Cambridge, MA: MIT Press/Bradford Books. (Cited by Allopenna et al., 1998, and Cutler, 1995)

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions during word-recognition in continuous speech. *Cognitive Psychology*, *10*, 29–63. (Cited by Allopenna et al., 1998, Tyler & Frauenfelder, 1987, and Cutler, 1995)

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: information processing time with and without saccades. *Perception & Psychophysics*,

53(4), 372–380. (Cited by Allopenna et al., 1998, Dahan et al., 2000, Tanenhaus et al., 2000, Weber & Cutler, 2004)

McClelland, J., & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86. (Cited by Allopenna et al., 1998)

Meľčuk, I. A. (1958). статистика и зависимость рода французских существительных от их окончания [Statistics and the relationship between the gender of French nouns and their endings]. *Бюллетень объединения по проблемам машинного перевода*, *7*, 13–40. (English version in Rozentsvejg, 1974, pp. 11–42. Cited on p. 57 by Corbett, 1991)

New, B., Pallier, C., Ferrand, L., & Matos, R. (2003). *Une base de données lexicales du français contemporain sur internet:* Lexique *2* [A database with French present-day lexical information on the internet: Lexique 2]. (Database Version 2.60. Original documentation published in 2001 in *L'Année Psychologique*(101), 447–462. Available from `http://www.lexique.org`)

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234. (Cited by Allopenna et al., 1998)

Norris, D., McQueen, J., & Cutler, A. (1995). Competition and segmentation in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 1209–1228. (Cited by Weber & Cutler, 2004)

Robert, P. (1995). *Le nouveau Petit Robert.* Paris: Dictionnaires Le Robert.

Rozentsvejg, V. J. (Ed.). (1974). *Essays on lexical semantics.* Stockholm: Scriptor. (Cited on p. 57 by Corbett, 1991)

Sabourin, L. (2003). *Grammatical gender and second language processing: An ERP study.* The Netherlands: University of Groningen. (PhD Dissertation)

Sachs, K., & Villatte, C. (1968). *Grand dictionnaire Langenscheidt, seconde partie: allemand-français.* Berlin-München: Langenscheidt-Larousse.

Schmidt, R. (1986). Was weiß der Artikel vom Hauptwort? Ein Beitrag zur Verarbeitung syntaktischer Beziehungen beim Lesen [How much does the article know about the noun? A contribution on the processing of syntactic relations during reading]. *Zeitschrift für experimentelle und angewandte Psychologie*, *23*, 150–163. (Cited by van Berkum, 1996, and Guillelmon & Grosjean, 2001)

Search Engine Showdown. (2002, December, 31). *Search engine showdown. The user's guide to web searching.* `http://www.searchengineshowdown.com/stats/sizeest.shtml`.

Sekerina, I. (2003). *Grammatical gender and mapping of referential expressions in russian.* Talk presented at the 9th Annual Conference on Architectures and Mechanisms for Language Processing, Glasgow, Scotland.

Spinelli, E., & Alario, F.-X. (2002). Gender context effects on homophone words. *Language and Cognitive Processes*, *17*(5), 457–469.

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3).

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, *29*(6), 557–590.

Tanenhaus, M. K., & Spivey-Knowlton, M. J. (1996). Eye-tracking. Special Issue of *Language and Cognitive Processes: A Guide to Spoken Word Recognition Paradigms*, *11*(6), 583–588.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*, 1632–1634.

Tucker, G. R., Lambert, W. E., & Rigault, A. A. (1977). *The French speaker's skill with grammatical gender: An example of rule-governed behavior.* The Hague: Mouton. (Cited on pp. 57–61 by Corbett, 1991)

Twain, M. (1880). *The Awful German Language.* (Appendix D from *A Tramp Abroad.* To be found on various websites: `http://www.google.com/search?q=%22the+awful+german+language%22`)

Tyler, L. K., & Frauenfelder, U. H. (1987). The process of spoken word recognition: An introduction. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition* (pp. 1–20). Cambridge, MA: MIT Press/Bradford Books.

Vroomen, J., & de Gelder, B. (1995). Metrical segmentation and lexical inhibition in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance*(21), 98–108. (Cited by Weber & Cutler, 2004)

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*, 1–25.

Zubin, D. A., & Köpcke, K.-M. (1981). Gender: A less than arbitrary grammatical category. In R. A. Hendrick, C. S. Masek, & M. F. Miller (Eds.), *Papers from the 7th regional meeting of the Chicago Linguistic Society* (pp. 439–449). Chicago: University of Chicago Press. (Cited on pp. 38–39 by van Berkum, 1996)

Zubin, D. A., & Köpcke, K.-M. (1986). Gender and folk taxonomy: The indexical relation between grammatical and lexical categorization. In C. Craig (Ed.), *Noun classes and categorization: Proceedings of a symposium on categorization and noun classification, Eugene, Oregon, October 1983.* Typological Studies in Language, 7 (pp. 139–180). Amsterdam: John Benjamins. (Cited on p. 321 by Corbett, 1991)

Zwitserlood, P. (1989). The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition*, *32*, 25-64. (Cited by Allopenna et al., 1998, and Dahan et al., 2000)