The Quarterly Journal of Experimental Psychology Section A

ISSN: 0272-4987 (Print) 1464-0740 (Online) Journal homepage: http://www.tandfonline.com/loi/pqja20

The Use of Stereotypical Gender Information in Constructing a Mental Model: Evidence from English and Spanish

Manuel Carreiras

To cite this article: Manuel Carreiras (1996) The Use of Stereotypical Gender Information in Constructing a Mental Model: Evidence from English and Spanish, The Quarterly Journal of Experimental Psychology Section A, 49:3, 639-663, DOI: 10.1080/713755647

To link to this article: http://dx.doi.org/10.1080/713755647

Published online: 22 Oct 2010.

Submit your article to this journal

Article views: 268

View related articles

Citing articles: 49 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=pqja20

Download by: [Stanford University] **Date:** 07 April 2017, At: 18:10

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 1996, *49A* (3), 639–663

The Use of Stereotypical Gender Information in Constructing a Mental Model: Evidence from English and Spanish

Manuel Carreiras
University of La Laguna, Tenerife, Spain
Alan Garnham, Jane Oakhill, Kate Cain
University of Sussex, Brighton, U.K.

Four experiments were carried out to investigate how general knowledge about the stereotypical gender of participants in a text influences comprehension. A self-paced reading task was used to present short texts comprising one, two, or three sentences. The first sentence of each text introduced a stereotypically masculine or feminine participant (e.g. doctor, nurse), or a neutral one. The last sentence introduced a pronoun (he/she) that could match or mismatch the gender of the referent. The first experiment, which was carried out in English, showed that reading times for the last sentence were longer when there was a mismatch than when there was a match between the gender of the pronoun in the last sentence and the stereotypical gender of the referent in the first sentence. In contrast to English, the gender of the participant can be disambiguated by a preceding article (*el/la*) in Spanish. The results of the second, third, and fourth experiments, which were carried out in Spanish, showed that reading times for the first sentences were longer when there was a mismatch than when there was a match between the gender of the article and the stereotypical gender of the participant. However, reading times for the last sentences did not differ. Overall, the results suggest that information about the stereotypical gender of the participants in a text is incorporated into the representation as soon as it becomes available, and that it affects the ease with which the text is understood.

Comprehension during reading involves processes ranging from recognition of certain patterns in the printed material, such as letters and words, to those that construct an abstract representation of the text. The construction of a representation of the situation

Requests for reprints should be sent to Alan Garnham, Laboratory of Experimental Psychology, University of Sussex, Brighton, BN1 9QG, U.K. E-mail alang@biols.susx.ac.uk

Our collaboration started with a NATO collaborative research grant CRG.890527 and is continuing with an "Acción Integrada" grant HB93–212 from the British Council and the Spanish Ministry of Education. This research was also supported by grants 91/103 from the Dirección General de Universidades del Gobierno Autónomo de Canarias and PB93–0562 from the DGICYT awarded to Manuel Carreiras, and grant RC00232439 "Mental Models and Interpretation of Anaphora" from the Economic and Social Science Research Council (UK) awarded to Alan Garnham and Jane Oakhill.

© 1996 The Experimental Psychology Society

described in the text—a mental model—is an incremental process. A mental model provides both a representation of what the text is about and at the same time a context for understanding the following sentence. So, as that sentence is read, there is an incremental updating of the representation on the basis of the present and past input. In other words, the resulting representation at any given moment guides the interpretation of subsequent input (Garnham, 1987; Glenberg, Meyer, & Lindem, 1987; Johnson-Laird, 1983; Sanford & Garrod, 1981; van Dijk & Kintsch, 1983).

The notion of a mental model also provides a useful theoretical framework for explaining how general knowledge is accessed and incorporated into the representation of the content of a text. The reason why mental models theory is a good framework for describing the role of general knowledge is that mental models are representations of bits of the world, and general knowledge is also knowledge about bits of the world, so there is no problem in fitting the two together. Moreover, "past input" does not just mean the previous sentence or the previous piece of the text. Readers are not "tabulae rasae" each time they face a new text. Rather, they draw on their knowledge of the world to make assumptions and inferences about the situation described in the text. In fact, many anaphoric expressions can *only* be interpreted correctly with the help of knowledge about the world. The antecedent for an anaphoric definite description may have to be determined by a bridging inference (Clark, 1977). For example, to understand:

John was late for work.
The bus had been caught in a traffic jam.

it is necessary to assume that the bus mentioned in the second sentence was John's means of transportation to work. For pronouns, world knowledge may be needed to decide between two or more potential antecedents; for example, in the sentence:

Sue sold her car to Pam because she had decided to take up cycling.

she refers to Sue. This assignment depends on knowledge about alternative means of transport, and the consequences of buying and selling.

A particularly interesting kind of inference that can be made on the basis of world knowledge is about the gender of one or more characters in a text, when their gender is not explicitly given. Mental models are representations of situations that contain people, animals, and objects. When one of these elements is introduced into a mental model, some of its features are encoded into the representation. For instance, it is assumed that when we encounter a proper name like "John", features such as male, animate, and so on are incorporated into the representation (Carreiras, Garnham, & Oakhill, 1993; Chang, 1980; Garnham & Oakhill, 1985). In the case of nouns that are gender-specific, such as "father" and "mother", or of proper names, such as "John" and "Pamela", that are conventionally male or female, information about the gender of the participants seems to be automatically included in the corresponding situation model. However, when the text does not provide explicit information about the gender of a participant, as in the case of many role names (e.g. secretary), prior knowledge in the form of a gender stereotype could be used to provide a default gender. So, if a character is referred to as "Mary" or "the nurse", a

reader could infer in both cases that this character is a female. The default gender is provided by the stereotypes present in our society, and this default seems to be assumed until proved wrong by other information in the text.

A popular riddle illustrates this idea. Its relevance to the issues discussed in this paper was pointed out by Sanford (1985: p. 311, 1987: p. 26, p. 34). The riddle is to make sense of the following apparently incongruent scenario:

> This morning a father and his son were driving along the highway to work, when they were involved in a horrible accident. The father was killed, and the son was quickly driven to the hospital, severely injured. When the boy was taken into the operating theatre, the surgeon exclaimed: "Oh, my God, this is my son!"

The incongruity appears only if we attribute gender values to the different characters. In the case of "father" and "son" the gender values are explicit, but we also assign "male" as a default gender to "surgeon", on the basis of social stereotypes.

The violation of gender expectations produced by role names is particularly interesting for a theory of discourse comprehension, as (in both English and Spanish) gender-marked pronouns are often used to reintroduce previously mentioned referents without repetition of the phrases originally used to introduce them. The mental models theory assumes that the linguistic "antecedent" of a pronoun has introduced a token representing an individual or a set of individuals into a mental model of the discourse (Garnham, 1987; Glenberg, Meyer, & Lindem, 1987; Johnson-Laird, 1983; van Dijk & Kintsch, 1983). So, assigning antecedents to pronouns involves accessing entities in a discourse model (but see Garnham, Oakhill, Ehrlich, & Carreiras, 1995). To interpret a pronoun appropriately, a reader must select an antecedent that matches the pronoun in gender (and number). When gender-specific pronouns (e.g. he or she) that are used to refer to characters can be resolved solely on the basis of their gender, they are usually understood more quickly than those that cannot be (Carreiras, Garnham, & Oakhill, 1993; Garnham & Oakhill, 1985; Garnham, Oakhill, Ehrlich, & Carreiras, 1995; Gernsbacher, 1989). However, characters are sometimes introduced into the discourse by role names (e.g. nurse). Most role names in English are not marked for gender and are subject to the influence of stereotypes. Therefore, if a person is introduced into a mental model using a role name, that person can subsequently be referred to by a masculine or a feminine pronoun, which can match or mismatch the stereotypical gender of the role name. An English noun phrase such as "the engineer" or "the nurse" gives no explicit information about the gender of its referent. However, knowledge about the world tells us that most engineers are men and most nurses are women. In addition, the surgeon riddle suggests that gender stereotype information is used immediately to encode gender. On this assumption, there is a knowledge-based inference from the fact that a particular person is a surgeon to the fact that a particular person is (probably) male. This inference is elaborative, as it is not necessary for establishing coherence at the point when it is made.

It has recently been proposed that discourse representations constructed by readers without specific goals or strategies are "minimal" (McKoon & Ratcliff, 1992). A minimal representation does not include all the inferences typically thought to be included in a

mental model. Instead, the only inferences constructed are those based on easily available knowledge and those needed for local cohesion. Elaborative inferences are, by definition, unnecessary for cohesion. So, for elaborative inferences the crucial question is whether the relevant knowledge is readily available. Only if it is will the inference be made automatically. However, the minimalist position provides no clear criteria for what knowledge is readily available. We cannot, for example, rely on intuition to tell us what information is readily available, in the sense intended in the minimalist account. This fact can be illustrated from a different case. McKoon and Ratcliff (1992) argued that knowledge about even highly probable instruments (e.g. a spoon to stir coffee) is not readily available, because these default instruments are apparently not inferred during reading.

The case of a character introduced into a discourse by a role name and then referred to by a pronoun can provide experimental evidence about the effect demonstrated by the riddle. Because role names in English do not always provide explicit information about gender, the gender of the character may have to be inferred using knowledge of the stereotypical gender of a profession. But if, as the text continues (e.g. when a pronoun is encountered), the gender of the character turns out to be different from the stereotypical gender, the mental model will need to be updated. The end of a sentence such as "the engineer . . . she" or "the nurse . . . he" should, therefore, be read more slowly than that of "the engineer . . . he" or "the nurse . . . she". A result of this kind would suggest that an inference about the gender of the antecedent was made when the role name was read. Unfortunately, we cannot tell in English whether the effect is partly early (an elaborative or forward inference) or wholly late (a backward inference when the pronoun is interpreted). However, linguistic differences between English and Spanish will help us to untangle the possible loci of this effect.

We will compare English passages, such as (1), in which the gender of the pronoun in the second sentence either matches (1b) or mismatches (1c) the stereotypical gender of the role name with corresponding Spanish passages, such as (2) and (3), in which the gender of the participants is explicitly instantiated by the article (*el/la*) preceding the role name.

(1a) The footballer wanted to play in the match.
(1b) HE had been training very hard during the week.
(1c) SHE had been training very hard during the week.

(2a) *El futbolista quería jugar el partido.*
(2b) *EL había estado entrenando mucho durante la semana.*

(3a) *La futbolista quería jugar el partido.*
(3b) *ELLA había estado entrenando mucho durante la semana.*

An early effect in Spanish, caused by a clash between explicit gender and stereotype, would suggest that the effect is early in English, too, with immediate activation of stereotype knowledge and the use of that knowledge to assign a (probable) gender to a person. Activation of stereotype knowledge *by itself* is not sufficient to cause a problem. There is no inconsistency between a particular engineer being female and most engineers

being male. A problem will arise only if the stereotype is used to suggest a gender *for the particular person mentioned in the passage*. Furthermore, in the Spanish sentences the definite article unambiguously assigns a gender to a person. It would be unexpected and inefficient if other types of information, such as stereotypes, were accessed with the specific purpose of "helping" to assign gender. Only if that information were activated for some other reason—to build a model of a person known to be an engineer or a secretary—might it be expected to interfere with gender assignment from a definite article.

Indeed, stereotype information in Spanish is not used to *infer* a person's gender, because gender is explicitly given by the definite article. If that information is used to make any inference about gender, it is the inference that this person does not fit the stereotype. However, the crucial fact about the Spanish data is that it can establish that stereotype information is used early (in English), because it is unlikely that Spanish nouns activate stereotype information but English nouns do not. In English, unlike Spanish, to assign gender early on the basis of stereotype information is to make a straightforward inference about gender. Therefore, if the predictions are met both in Spanish and English, from the English and Spanish data together we get evidence for an effect in English that is both early and inferential, and which is also probably *nonminimal*.

EXPERIMENT 1

The purpose of the first experiment was to investigate the role of stereotype information in anaphor-antecedent mapping, using sentences like (1), in which the antecedents are professions that are either typically male, typically female, or neutral (i.e. free from gender bias). Many role names (i.e. occupational titles), including all those in our experimental materials, are not marked for gender in English and are subject to the influence of stereotypes. When reading sentences such as (1a), it is likely that we incorporate the stereotypical gender of the profession into the representation of the particular person mentioned in the text. If so, it should take subjects longer to read the second sentence when there is a mismatch between the gender of the pronoun and the stereotypical gender of the role name than when there is a match.

Method

Subjects

The subjects were 24 staff and students from Sussex University, and they were paid for taking part.

Materials

In order to select professions that were biased either to masculine or to feminine stereotypes or were neutral, 30 subjects, who did not take part in the main reading experiment, were asked to rate 120 professions for "gender stereotypicality": the likelihood that each would be done by either a man or a woman. It was stressed to subjects that their responses should reflect the way the world is, not how they might feel it should be. In the on-line reading experiment we used the 20 best exemplars

from each of the three categories (male, female, and neutral) chosen from the rating study. The English role names selected are presented in Appendix 1, with their ratings.

For each profession, a two-line text was written. Altogether, 60 texts were constructed (20 for male-biased characters, 20 for female-biased characters, and 20 for neutral characters). The subject of the first sentence was always an individual introduced only by a job title that was stereotypically male (e.g. 4a) or female (e.g. 4c) or was not stereotyped for gender (e.g. 4e). The second sentence (e.g. 4b, 4d, 4f) contained an anaphoric reference to this individual, which either matched or mismatched the gender stereotype in the case of stereotyped characters. However, there could be no mismatch between the gender of pronoun and the gender of the stereotype when the role name was neutral.

(4a) The electrician examined the light fitting.
(4b) He/She needed a special attachment to fix it.
(4c) The baby-sitter settled down to watch a video.
(4d) Then he/she heard the baby crying.
(4e) The psychologist studied the students' test results.
(4f) He/She was concerned about the low standard.

For each passage, a simple yes/no question was constructed. For example, the question for the passage in (4a) and (4b) was: "Was the electrician mending a stereo?"

Procedure

Subjects were tested individually in a small, quiet room. Each subject read 60 experimental passages. There were 20 items from each category (male, female, and neutral stereotypes), 10 in which the pronoun was masculine, and 10 in which it was feminine. There were also 24 filler items. Half of these contained stereotypically male professions, and half contained stereotypically female professions. These were followed by their gender match pronoun. The fillers were included to reduce the proportion of texts about people whose occupations did not match the gender stereotype, so that subjects would not realize that these texts were under investigation.

The experiment was controlled by a custom-built 6809-based microprocessor. The two sentences of each passage and the question were presented successively in the centre of the screen. A two-button box was located between the subjects and the screen, with one button labelled "yes" and the other labelled "no".

Before each text the prompt, NEXT TEXT, appeared on the screen. When it was presented, subjects had to press the button on the side of their dominant hand to display the first sentence. Another press displayed the second sentence. A third press displayed the question. Subjects pressed the button under their dominant hand to answer "yes". They pressed the other button to answer "no". The instructions stressed that the sentences were to be read at normal reading speed, but that the questions were to be answered as quickly and as accurately as possible.

Before the presentation of the experimental materials, there were six practice trials, whose primary purpose was to familiarize subjects with the self-paced reading procedure.

Results and Discussion

Table 1 shows the mean reading times for the second clauses of sentences that contained stereotyped and neutral characters. All the data were subjected to two analyses, one treating subjects as a random factor and one treating items as a random factor. For the sentences that contained stereotyped characters, 2×2 ANOVAs were performed with the

TABLE 1
Mean Reading Times for the Second Sentences in Experiment 1

	Stereotyped Characters		*Neutral Characters*	
	Match	*Mismatch*	*Masculine*	*Feminine*
male bias	1651	1807		
female bias	1632	1735		
TOTAL	1641	1771	1708	1708

Note: Stereotypically male and female characters in the match and mismatch conditions; characters that were not stereotyped for gender preceded by a masculine or a feminine article.

factors: stereotypical gender of the character, and match or mismatch of the pronoun's gender with the stereotype. We also compared reading times for the second clauses of sentences that contained neutral characters, followed by a masculine or feminine pronoun.

The ANOVAs revealed that only the main effect of gender matching was reliable, $F1(1, 23) = 28.31$, $p < .0001$; $F2(1, 38) = 7.97$, $p = .01$. Second-sentence reading times were slower in the mismatch condition than in the match condition. However, neither the main effect of stereotypical gender of the character, $F1(1, 23) = 2.37$; $F2(1, 38) = 0.37$, nor the interaction, $F1(1, 23) = 1.50$; $F2(1, 38) = 0.30$, was reliable.

There were no significant reading-time differences between neutral professions followed by a masculine pronoun and those followed by a feminine pronoun, $F1(1, 23) = 0.00$; $F2(1, 19) = 0.00$. For the neutral professions, target sentences with a masculine pronoun were read as fast as sentences in which the pronoun was feminine.

In order to make a direct comparison between the stereotyped and neutral characters, we performed an analysis with character type (male stereotype, female stereotype, neutral) and gender of pronoun (masculine, feminine) as factors. We coded the factors in this way to avoid an arbitrary classification of masculine and feminine pronouns as matching or mismatching the neutral characters. In this analysis, there was a Character Type $\times$ Gender of Pronoun interaction, $F1(2, 46) = 11.15$, $p < .0001$; $F2(2, 57) = 4.81$, $p < .05$. This interaction arises because there was no difference in the reading times for the sentences containing the male and female pronouns when they referred to the neutral characters, but there were matching effects for the male and female stereotypes.

One question that the analyses reported so far do not answer is whether subjects were slowed down by a mismatch between the stereotype and the gender of the pronoun, speeded up by a match, or both. This question cannot be answered from the overall reading times, because the materials containing the stereotypes were different from those containing the neutral characters, so we would be looking for a main effect on a between-items comparison. Any differences in the overall reading times might reflect differences in, for example, the lengths of the sentences. In order to correct for this problem, we performed an analysis on the reading times per character for the second sentences, in which we compared the stereotype-match, stereotype-mismatch, and the neutral conditions. This analysis can only be carried out straightforwardly by subjects, as the same 40 items contribute to the first two conditions, but a separate set of 20 items contributes to

the third. The mean per-character reading times in the three conditions were 44.2 msec, 48.0 msec, and 42.9 msec. There was a main effect of condition, $F1(2, 46) = 25.89$, $p < .001$. The stereotype-mismatch condition was significantly slower than the neutral condition, $F1(1, 23) = 38.58$, $p < .001$. However, the difference between the stereotype-match condition and the neutral condition just missed significance, $F1(1, 23) = 4.23$, $0.1 > p > .05$. These tests are not independent. Dunnett's post hoc test, with the neutral condition regarded as the control, showed that only the stereotype mismatch condition differed from the neutral, $p < .01$, one-tailed. In any case, the neutral condition was faster than the stereotype-match condition, so there was no evidence that a match between the gender of the pronoun and the gender of the stereotype helped in mapping the pronoun onto its referent. Matching to a stereotype is no easier, and no harder, than assigning a gender from a pronoun to a character with an unbiased role name. On the other hand, a mismatch with a stereotype slowed people down considerably.

It might appear that the lack of a difference between the match and the neutral conditions suggests that an inference about gender has not been made in the match condition. However, even if an inference has been made, it does not follow that the match condition should be faster. In the neutral condition, the pronoun establishes the gender and allows it to be encoded into the model. In the match condition, if a forward inference has been made, the gender of the pronoun has to be mapped onto the already established gender of the person bearing the role name. However, we know from the mismatch condition that stereotype information is activated when the pronoun is read. So, in the match condition there will be an extra process of checking to see whether the gender of the pronoun matches that of the stereotype. As we have no a priori reason to believe that establishing gender from a pronoun, in the absence of other information (neutral condition), is easier or more difficult than mapping to a previously established gender and checking for consistency with a stereotype (match condition), we cannot make any clear prediction about the relation between the match and neutral conditions.

There are two possible explanations of the match/mismatch effect in the sentence that contains the pronoun. The first is that the stereotype information is activated when the NP (e.g. the electrician) is initially encoded and used to decide that this particular electrician is probably male. On this view, the match/mismatch effect arises when the pronoun is linked to the person whose (probable) sex is already encoded. If the pronoun disconfirms the gender indicated by the stereotype, the mental model has to be updated, resulting in longer reading times for those sentences.

The other explanation is that the stereotype information is not activated by the encoding of the NP, but only when the morphologically marked pronoun has to be linked to it. In this case, the information about the stereotyped gender of the protagonist is not included in the mental model. We regard this explanation as unlikely, for the same reason that we regard it as unlikely that stereotype information is recruited specifically by processes that assign gender from morphological markings on definite articles in Spanish. Assuming that the pronoun is taken to refer to the character introduced by the role name, the gender of the pronoun provides clear information about the gender of that person. To activate other less reliable and possibly conflicting information at this point to help assign gender is inefficient. However, when the role name is first read, there is no guarantee that more reliable information about gender will follow. It can, therefore, be argued that the

stereotype information is more useful at that point. However, this argument is weak, compared to the one we gave for Spanish, because, in general, a pronoun will have several potential referents, and stereotype information might be useful, at the point of pronoun resolution, in choosing between them.

In our second experiment, we explored the two alternative hypotheses about when stereotype information is brought to bear on the assignment of genders to characters. The gender of a Spanish noun can be unambiguously signalled by an accompanying definite article. Thus, unlike in the case of English pronouns, where stereotype information may help in deciding which person's gender is being specified, the use of stereotype information can only complicate the process of gender assignment in these Spanish cases. If Spanish readers use stereotype information, it is almost certainly not because they have a procedure that activates this information when they have to work out the gender of the referent of a noun phrase from that noun phrase's morphology. More probably, it comes from processes that enter tokens representing referents with certain properties (e.g. being an engineer) into mental models. And if Spanish readers use stereotype information in this way, it is almost certain that English readers do, too.

On this view, the gender of the person, as indicated by the definite article, should be incorporated into the mental model, together with the role name. However, stereotype information is also accessed *via the role name*, when that name is used to construct a representation of the person in a mental model. We should, therefore, expect longer reading times when there is a mismatch rather than a match between the morphological gender of the article and the stereotypical gender of the role name. However, we should not expect any differences to appear in the sentence with the pronoun, as the mismatch between the stereotypical gender and the actual gender will already have been resolved. On the other hand, if the English findings are explained by backwards inference, then in Spanish gender will be encoded in the first sentence from the morphological information on the definite article, but stereotype information should not be activated. Therefore, we should expect a late effect in Spanish, as in English—one that occurs when the stereotype information is activated on linking the pronoun.

EXPERIMENT 2

Method

Subjects

The subjects were 48 undergraduate students from the University of La Laguna, who participated in the experiment for course credit.

Materials and Design

As in Experiment 1, we collected ratings on stereotypes for professions in Spanish, in order to select professions with male and female stereotypes, and neutral professions, for the reading time study. Sixteen subjects, none of whom took part in the reading experiment, were asked to rate 200 professions for "gender stereotypicality". Subjects were asked to rate each profession for the like-

lihood that it would be done by either a man or a woman. It was stressed to subjects that their responses should reflect the way the world is, not how they might feel it should be.

In the on-line reading experiment we used the 12 best exemplars from each of the three categories (male, female, and neutral) chosen from the rating study. The Spanish role names selected, together with their ratings, are included in Appendix 2. For each profession, a short text was written. Altogether 36 texts were constructed (12 with male biased characters, 12 with female biased characters and 12 with neutral characters), each with three sentences. The first sentence of each text introduced a character, using a role name that was either stereotypically male or female or not stereotyped for gender. The name was preceded by a masculine or feminine definite article (*el/la*). The second sentence contained information consistent with the first sentence. This sentence was included to catch any spillover effects from the first sentence, so we could be sure that if there was an effect of match/mismatch in the pronoun sentence, it was a real one. Because of the hypotheses we were testing, it was essential that we were able to locate the mismatch effect in Spanish accurately. In English, the locus of the effect is inevitably ambiguous, so we did not need to worry about spillover of stereotype-based processing, from the first sentence to the second. The third sentence began with a masculine or a feminine pronoun (*el/ella*), which referred to the character mentioned in the first sentence.

Stereotypically masculine

El carpintero/La carpintera tomó las medidas para hacer el armario.
Era un encargo bastante urgente.
El/Ella tenía que terminarlo en el plazo de una semana.

The carpenter took measurements to make the cupboard.
It was a quite urgent order.
He/She had to finish it in the space of one week.

Stereotypically feminine

La enfermera/El enfermero tuvo que suturar la herida.
El corte había sido profundo.
Ella/El también puso una inyección para evitar una infección.

The nurse had to suture the injury.
It had been a deep cut.
She/He also gave an injection to avoid an infection.

Not stereotyped for gender

El cantante/La cantante recibió una gran ovación.
El ambiente de fiesta era impresionante en el estadio.
El/Ella tuvo que salir de nuevo al escenario.

The singer received a great ovation.
The festival atmosphere was impressive in the stadium.
He/She had to appear on the stage again.

The 24 texts about stereotyped characters (12 male, 12 female) were constructed as follows. The first sentences combined two factors: (1) whether the stereotypical gender of the character was male or female, and (2) gender matching—whether the gender of the definite article matched or mis-

matched the stereotypical gender of the character. In addition, the third sentence combined two other factors: (1) morphological matching—whether the initial pronoun of the third sentence matched or mismatched the gender of the definite article in the first sentence—and (2) stereotype matching—whether the gender of the pronoun matched or mismatched the stereotypical gender of the character in the first sentence.

In the passages with characters that were not stereotyped for gender, the character introduced in the first sentence was preceded by a masculine or a feminine article. In addition, the initial pronoun of the third sentence matched or mismatched the gender of the article that appeared in the first sentence.

Procedure

The procedure was similar to that of Experiment 1, except that an IBM-compatible computer was used, and subjects pressed the space bar to advance the display, the "L" key to answer "yes", and the "A" key to answer "no". The Spanish labels for "yes" and "no" were fixed on those keys. The prompt before each passage was COMIENZA OTRO TEXTO [another text is ready]. Each subject read 36 experimental texts, intermixed randomly with 128 others that served as fillers. In this and the following experiments, the fillers included other texts containing gender-marked pronouns, with either animate or inanimate antecedents, structurally ambiguous sentences, and a variety of other sentences designed to distract subjects' attention from the form of the experimental items. Some of the filler sentences were followed by questions, but the experimental sentences were not. As in Experiment 1, there were 6 practice items.

Results and Discussion

Table 2 shows the mean reading times for the first and second sentences of the passages in Experiment 2.

In the main analyses, data for sentences containing male and female stereotypes were pooled. This pooling was justified by the results of initial 2 × 2 ANOVAs, performed by

TABLE 2
Mean Reading Times for the First and Second Sentences in Experiment 2

		Stereotyped Characters		*Neutral Characters*	
Clause		*Match*	*Mismatch*	*Masculine*	*Feminine*
1	male bias	3193	3400		
	female bias	3191	3454		
	TOTAL	3192	3427	3121	2992
2	male bias	2213	2211		
	female bias	2282	2348		
	TOTAL	2248	2280	2621	2504

Note: Stereotypically male and female characters in the match and mismatch conditions; characters that were not stereotyped for gender preceded by a masculine or a feminine article.

subjects and by items, with stereotypical gender of the character (male vs. female) and gender matching (match vs. mismatch) as factors. Reading times for the first clauses did not show any differences for stereotypical gender of the character, $F1(1, 47) = 0.12$; $F2(1, 22) = 0.02$; neither was there an interaction between stereotypical gender of the character and gender matching, $F1(1, 47) = 0.11$; $F2(1, 22) = 0.07$.

In the combined analysis, first clause reading times were slower in the mismatch condition than in the match condition, $F1(1, 47) = 8.43$, $p < .01$; $F2(1, 23) = 4.98$, $p < .05$. It took longer to read the first sentences when there was a mismatch than when there was a match between the gender of the article and the stereotypical gender of the character that followed. Reading times of second clauses in the match and mismatch conditions did not show reliable differences, $F1(1, 47) = 0.26$; $F2(1, 23) = 0.17$.

Separate ANOVAs were also performed on the mean reading times for the first and second sentences of the passages with characters that were not stereotyped for gender. The ANOVAs showed no reliable effect of the gender of the definite article either for the first sentence, $F1(1, 47) = 1.54$; $F2(1, 11) = 0.59$, or for the second sentence, $F1(1, 47) = 1.34$; $F2(1, 11) = 0.82$.

In order to make a direct comparison between stereotyped and neutral role names, we performed an analysis with character type (male stereotype, female stereotype, neutral) and gender of definite article (masculine, feminine) as factors. In this analysis, there was a Character Type $\times$ Gender of Article interaction, $F1(2, 94) = 4.72$, $p < .01$; $F2(2, 33) = 2.35$. This interaction arises because there was no difference in the reading times for the sentences with masculine and feminine definite articles when the role names were neutral, but there were matching effects for the male and female stereotypes.

The by-subjects analysis on the reading times per character for the first sentences, in which we compared the stereotype-match, stereotype-mismatch, and the neutral conditions, showed a main effect of condition, $F1(2, 94) = 5.74$, $p < .005$. The stereotype-mismatch condition was significantly slower than the neutral condition, $F1(1, 47) = 7.26$, $p < .01$. However, there was no difference between the stereotype-match condition and the neutral condition, $F1(1, 47) = 0.41$. The mean reading times per character for the first sentences in the three conditions were: stereotype match, 64.1 msec; stereotype mismatch, 69.1 msec; and neutral, 65.1 msec. Dunnett's test showed that only the stereotype mismatch condition differed from the neutral condition, $p < .01$, one-tailed. The neutral condition was closer to the stereotype-match condition than to the stereotype mismatch, so as in Experiment 1 there was no evidence that a match between the morphologically specified gender and the gender of the stereotype speeded subjects' reading. This fact is consistent with our claim that the lack of difference between the match and neutral conditions in Experiment 1 cannot be used as evidence against early use of stereotype information to encode gender.

Table 3 shows the mean reading times for the third sentences. A 2×2 ANOVA for the passages with stereotyped characters, with morphosyntactic matching and stereotype matching as factors, showed that only the main effect of morphosyntactic matching was reliable, $F1(1, 47) = 44.24$, $p < .0001$; $F2(1, 23) = 16.12$, $p < .0005$. Neither the main effect of stereotype matching, $F1(1, 47) = 0.08$; $F2(1, 23) = 0.07$, nor the interaction between the two factors, $F1(1, 47) = 0.39$; $F2(1, 23) = 0.35$, was reliable. In addition, a separate ANOVA on the mean reading times for the third sentences in the neutral

TABLE 3
Mean Reading Times[a] for the Third Sentences in Experiment 2

	Morphosyntactic	
	Match	*Mismatch*
Stereotype match	2374	2665
Stereotype mismatch	2356	2723
TOTAL	2365 (55)[a]	2694 (63)
Neutral characters	2736 (56)	2892 (60)

[a] Msec. per character in parentheses.

Note: Stereotypically male and female characters in the match and mismatch conditions; characters that were not stereotyped for gender preceded by a masculine or a feminine article.

condition showed that the tendency to be slower in the mismatch than in the match condition did not reach statistical significance, $F1(1, 47) = 3.16$, $p < .09$; $F2(1, 11) = 1.67$.

Finally, to compare directly the morphoysyntactic matching effects for sentences that contained stereotyped role names and neutral names, we performed 2 × 2 ANOVAs by subjects and items with linguistic match and type of name (stereotyped vs. neutral) as factors. To correct the problem of having sentences with different lengths, the anlayses were performed on the per-character reading times for the third sentences, which are also shown in Table 3. The ANOVAs were performed with equally weighted means: Note that a different number of observations contributes to the subject-by-condition means in the by-subjects analysis, and that there are different numbers of items of the two types in the by-item analysis. As expected from the previous analyses, the ANOVAs showed a main effect of linguistic match, $F1(1, 47) = 23.35$, $p < .0001$; $F2(1, 34) = 11.71$, $p < .005$. They also showed an interaction between the two factors, but it was only significant in the by-subjects analysis, $F1(1, 47) = 9.33$, $p < .005$; $F2(1, 34) = 2.36$. The third sentences were read faster when there was a match between the gender of the article in the first sentence and the gender of the pronoun in the third sentence, both when the first sentences contained stereotyped role names and when they contained neutral role names. This interaction arises because reading is slowed down more by a morphological mismatch when the characters are stereotyped. There was little difference between the reading times for stereotyped and neutral characters when the morphological gender of the pronoun matched that of the definite article.

The results of Experiment 2 suggest that world knowledge about social stereotypes is activated at the earliest possible point, which in these Spanish texts is when the character's name (role name or occupational title) is first mentioned. The token representing the character in the mental model includes information about the person's gender (from the definite article). When the default gender provided by the social stereotype disagrees with the morphological gender of the definite article text, processing suffers, resulting in longer reading times than when the two genders agree. As with the corresponding (but later) effect in the English experiment, this effect was attributable to a slowing-down in

the mismatch condition as compared with the neutral condition, and not to a speeding-up in the match condition.

It is important to notice that once the actual gender is incorporated into the discourse model, the default gender provided by the stereotype does not influence the time it takes to interpret a subsequent pronoun. In contrast to English, when the gender of the pronoun is in agreement with the morphological gender of the definite article in the first sentence, text processing does not suffer, even though the gender of the pronoun does not match the stereotypical gender of the character mentioned in the first sentence. The only reliable effect in the reading times of the third sentences was that of morphosyntactic matching. There was no effect of stereotype matching. It seems that the gender of the character mentioned in the first sentence is determined by the definite article and incorporated into the representation when that is read. A mismatch with the stereotypical gender causes problems at that point, but they are resolved by the end of the first sentence. Thus, the later effect of morphosyntactic matching is independent of stereotype matching. An alternative possibility, though one that this experiment rules out, would have been that two mismatches (*la futbolista . . . el . . .*, "the -feminine- footballer . . . he . . .) partly cancelled each other out because the masculine implication of footballer was not fully suppressed.

The morphosyntactic matching effect does, however, interact with whether the characters are stereotyped or neutral, with the difference arising in the morphosyntactic mismatch condition. A plausible explanation of this effect is that when subjects detect a morphological mismatch, they know that there is something wrong with the text or their reading of it, and they recruit whatever information is available in an attempt to solve their problems. However, stereotype information, whether it matches or mismatches the gender of the original definite article, cannot resolve the conflict between the morphology of the pronoun and the morphology of the article. It does, however, provide an extra piece of information to take into consideration, and that, presumably, is why subjects slow down.

It was important to include both morphosyntactic and stereotype matching in the third sentences of Experiment 2 to isolate the effect of each factor. However, the morphological mismatch between the gender of the definite article and the gender of the pronoun in some of the passages might induce subjects to develop a special strategy. It is quite striking for a Spanish reader to find a masculine pronoun used to refer to a character previously introduced by a noun phrase with a feminine article or a feminine pronoun referring to a character previously introduced by a noun phrase with a masculine article. In order to rule out the possibility that the stereotype mismatch effect was affected by special strategies, we carried out another experiment in which these particular combinations of definite articles and pronouns were not included.

EXPERIMENT 3

Method

Subjects

The subjects were 32 undergraduate students from the University of La Laguna, who participated in the experiment for course credit.

Materials and Design

The materials were the same as in Experiment 2, with two exceptions. First, the initial pronoun of the third sentence always matched the gender of the definite article that preceded the character name in the first sentence. Second, a Yes–No question was constructed for each of the sentences to encourage subjects to read the sentences carefully and to understand them. To answer this question, subjects had to resolve the pronoun in the third clause. Half of the questions had ''yes'' as their answer and the rest ''no''. The questions corresponding to the sentences given for Experiment 1 were:

¿Tenía el carpintero una semana de plazo?
Did the carpenter have a week (to do the job)?

¿Puso la enfermera una inyección?
Did the nurse give an injection?

¿Tuvo que salir la cantante de nuevo al escenario?
Did the singer have to appear on the stage again?

Procedure

The procedure was similar to that of Experiment 2. Each subject had to read the 36 experimental sentences intermixed randomly with each other, and with 83 other sentences, which acted as fillers. Some of the filler sentences were also followed by questions.

Results and Discussion

Table 4 shows the mean reading times for the three sentences, the percentage of questions correctly answered, and the question answering times for Experiment 3. We performed separate ANOVAs on the reading times for each sentence, on the percentage of questions correctly answered, and on the question answering times, both for stereotyped role names and for neutral names. The ANOVAs for stereotyped role names showed that the only reliable difference between the match and mismatch conditions appeared in the first sentence. Reading times for the first sentence were slower in the mismatch than in the match condition, $F1(1, 31) = 11.61$, $p < .005$; $F2(1, 23) = 7.55$, $p < .05$. However, no other differences were reliable: reading times for the second sentence, $F1(1, 31) = 1.01$; $F2(1, 23) = 0.77$, reading times for the third sentence, $F1(1, 31) = 0.12$; $F2(1, 23) = 0.05$, percentage of questions correctly answered, $F1(1, 31) = 0.35$; $F2(1, 23) = 0.22$, question answering times, $F1(1, 31) = 2.59$; $F2(1, 23) = 1.57$.

TABLE 4
Mean Reading Times for the First, Second, and Third Sentences, Percentage of Questions Correctly Answered, and Question Answering Times in Experiment 3

	Stereotyped Characters		*Neutral Characters*	
	Match	*Mismatch*	*Masculine*	*Feminine*
Clause 1	2929	3154	2793	2837
Clause 2	2092	2051	2332	2339
Clause 3	2141	2151	2450	2447
Percentage of correct responses	95	96	98	96
Question answering times	2399	2506	2805	2761

Note: Stereotypically male and female characters in the match and mismatch conditions; characters that were not stereotyped for gender preceded by a masculine or a feminine article.

The ANOVAs for the sentences that contained a neutral role name did not show any reliable difference between masculine and feminine articles and pronouns: first sentence, $F1(1, 31) = 0.26$; $F2(1, 11) = 0.24$, second sentence, $F1(1, 31) = 0.11$; $F2(1, 11) = 0.01$, third sentence, $F1(1, 31) = 0.00$; $F2(1, 11) = 0.00$, percentage of hits, $F1(1, 31) = 1.35$; $F2(1, 11) = 0.74$, question answering times, $F1(1, 31) = 0.27$; $F2(1, 11) = 0.14$.

To make a direct comparison between the stereotyped and neutral role names, we performed an analysis with character type (male stereotype, female stereotype, neutral) and gender of definite article (masculine, feminine) as factors. In this analysis there was a Character Type $\times$ Gender of Article interaction, $F1(2, 62) = 7.19$, $p < .005$; $F2(2, 33) = 4.24$, $p < .05$. This interaction arises because there is no difference in the reading times for the sentences with masculine and feminine definite articles when the role names are neutral, but there are matching effects for the male and female stereotypes.

As in the previous experiments, we also performed an analysis on the reading times per character for the first sentences, in which we compared the reading times for the stereotype-match, stereotype-mismatch, and the neutral conditions. The mean per-character reading times in the three conditions were 58.8 msec, 63.6 msec, and 59.8 msec. There was a main effect of condition, $F1(2, 62) = 6.48$, $p < .005$. The stereotype-mismatch condition was significantly slower than the neutral condition, $F1(1, 31) = 5.80$, $p < .05$. However, the difference between the stereotype-match condition and the neutral condition was not reliable, $F1(1, 31) = 0.57$. Dunnett's post hoc test confirmed that the difference between the neutral and stereotype mismatch conditions was reliable, $p < .01$, one-tailed, but that the difference between the neutral and stereotype match conditions was not.

Again, the results of this experiment showed reliable differences in the reading times of the first sentences only. Reading times were longer when the default stereotypical gender of the role names and the gender of the definite article mismatched than when they matched. So the results of Experiment 3 replicate those of Experiment 2. The stereotypical gender associated with the role names affects the initial encoding of the characters' gender, which is determined by the morphological properties of the noun phrase. But once the gender is incorporated into the discourse model, the default gender provided by the social stereotype does not influence the times taken to resolve a subsequent pronoun.

The whole sentence reading times that we measured in Experiments 2 and 3 can only give a crude idea of the locus of the effect of stereotype information. We cannot yet conclude that the article–noun combination is sufficient by itself for the effect, which may arise only when that article–noun combination is interpreted in its sentential context. If this idea is correct, the effect should show up only at the end of the sentence. If not, then we expect an effect when only the NP has been read. This hypothesis can be tested by considering just the first sentences and presenting them in two displays.

Experiment 4 was designed to distinguish between these hypotheses. The first sentences of the passages from Experiments 2 and 3 were presented, with the article and the noun in one display and the rest of the sentence in a second display. If the incongruence between the morphological gender of the definite article and the stereotype is sufficient to produce the effect, reading times for the first display should be longer in the mismatch condition than in the match condition. As we found no effects on the reading times for the first sentences in the neutral conditions in Experiments 2 and 3 and no differences between the neutral and match conditions, we did not include the neutral condition in Experiment 4.

EXPERIMENT 4

Method

Subjects

The subjects were 72 undergraduate students of the University of La Laguna, who participated in the experiment for course credit.

Materials and Design

Only the first sentences of the previous experiments that contained a stereotypically male or female character, preceded by a masculine or feminine definite article (*el/la*), were used: 12 with stereotypically masculine characters and 12 with stereotypically feminine characters. The sentences were constructed by combining two factors: (1) whether the stereotypical gender of the character was masculine or feminine, and (2) gender matching—i.e. whether the gender of the article matched or mismatched the stereotypical gender of the character.

The sentences were divided into two displays. The first contained the article followed by the role name, and the second included the rest of the sentence. The slashes signal the point where the sentence was segmented into two displays.

(El carpintero) (La carpintera)/tomó las medidas para hacer el armario.
The carpenter/took measurements to make the cupboard.

(La enfermera) (El enfermero)/tuvo que suturar la herida.
The nurse/had to suture the injury.

Procedure

The procedure was similar to that of Experiment 3, except that there was only one sentence, and it was divided into two displays that were presented successively in the centre of the screen. Each subject read the 24 experimental sentences intermixed randomly with 142 other sentences that served as fillers. Some of the filler sentences were followed by questions, but the experimental sentences were not. Again there were 6 practice items before the main part of the experiment.

Results and Discussion

Table 5 shows the mean reading times for the two displays of the sentences. We performed 2 ANOVAs (by subjects and by items) with stereotypical gender of the characters and gender matching as factors, separately for each display. First display reading times were slower in the mismatch condition than in the match condition, $F1(1, 71) = 49.89$, $p < .0001$; $F2(1, 22) = 33.82$, $p < .0001$. The interaction between the two factors was also reliable, $F1(1, 71) = 4.10$, $p < .05$; $F2(1, 22) = 9.98$, $p < .005$. First display reading times of sentences that contained stereotypically masculine characters were slower in the mismatch condition than in the match condition, $F1(1, 71) = 68.58$, $p < .0001$; $F2(1, 11) = 43.41$, $p < .0001$. However, the reading time differences between match and mismatch conditions in the first display of sentences that contained stereotypically feminine characters were only significant by subjects, $F1(1, 71) = 7.78$, $p < .01$; $F2(1, 11) = 3.16$. However, because of the lack of power by items, we would not conclude that there is no effect with female stereotype role names.

The ANOVAs on the second display showed that reading times were slower in the mismatch condition than in the match condition, but only by subjects, $F1(1, 71) = 9.17$, $p < .01$; $F2(1, 22) = 0.48$. The interaction between this factor and whether the character was male or female was not reliable, $F1(1, 71) = 1.70$; $F2(1, 22) = 0.00$.

A further ANOVA, with display as a factor, showed an interaction between display and matching that was significant by items, but not by subjects, $F1(1, 71) = 0.81$; $F2(1, 23) = 4.99$, $p < .05$.

TABLE 5
Mean Reading Times for the First and Second Displays of Sentences with Stereotypically Male and Female Characters in the Match and Mismatch Conditions of Experiment 4

Display		*Match*	*Mismatch*
1	male bias	906	1019
	female bias	943	1003
	TOTAL	924	1011
2	male bias	1644	1736
	female bias	1694	1738
	TOTAL	1669	1737

The results indicate that although the scenario created by the whole sentence might contribute to the effects found in the previous experiments, the effect of the stereotypical gender of the role names occurs when only the role name and the definite article are presented. Indeed, the effect is stronger in the first frame, and it is particularly weak by items in the second frame. One possibility is that the context only reinforces the stereotyping significantly in a few of the passages. Another is that the effect in the second frame is a spillover effect.

The effect of matching is weaker for female stereotypes, which can be explained by the fact that the Spanish female stereotypes were weaker in the rating data than the male ones (see Appendix 2). The examples used in the experiment were those that produced the most biased results in the norming study. It was more difficult to find role names with a female bias than role names with a male bias, and it was especially difficult to find role names with a very strong female bias.

GENERAL DISCUSSION

Taken together, the results of the four experiments make a strong case for the view that stereotypical gender is incorporated into the representation of characters introduced by role names and updated by incoming information. When a role name such as "footballer" is read, a representation of a person who is a footballer is inserted into the mental model being developed. If, as in English, there is no morphological marker for gender in the corresponding noun phrase ("the footballer") and no other explicit information in the text, the person is assumed to be male because the stereotypical gender of footballers is male. If a pronoun encountered later in the text has the same gender as the person represented in the mental model, referential assignment is easy, though the results of Experiment 1 suggest that the assignment is no easier than in the case where no stereotyping has occurred, and the gender of the pronoun determines, for the first time, the gender of the character (e.g. the student . . . he). However, if the pronoun is inconsistent with the stereotypical gender, the model must be updated, and that updating takes time. In this mismatch condition some additional processing is required for the pronoun.

The English data alone do not force the conclusion that an inference is made when a representation of a character of the type denoted by a role name is encoded into the mental model. They are consistent with early elaborative inferencing, which is not necessary for local cohesion, and which may not be based on readily available knowledge. But they are also consistent with the alternative idea that all the inferential work is done when the pronoun is read.

In Spanish, at least where there are no morphosyntactic inconsistences such as those we deliberately introduced in Experiment 2, the gender of the person introduced by a role name does not have to be updated in the mental model. The reason is that, when "*futbolista*" is encountered in a text, the default gender is compared with, and if necessary overridden by, the gender indicated by the local morphology: the form of the article, and in some cases the morphologically feminine or masculine ending of the noun. The Spanish data suggest that the stereotype information is activated when the information carried by the NP is encoded into the mental model, and that it affects the assignment of gender to the referent of the NP. Otherwise, readers would simply encode the fact that

someone was male or female from the morphological information: Stereotype information can have no effect unless it is specifically brought to bear on the gender of the individual the passage is about. Therefore, as the Spanish data establish that the effect of stereotyping is early and it is unlikely that English and Spanish speakers differ in when stereotype information is brought to bear on the interpretation of a sentence containing a stereotyped role name, it can be concluded that elaborative, and probably non-minimal, inferences are made in English.

One argument against this idea is that the stereotype information is brought to bear on the interpretation of the Spanish sentences only because the morphological marking on the definite article forces a gender assignment to be made. Although we cannot conclusively prove that this idea is incorrect, it commits its proponents to an implausible view of why stereotype information is activated. The morphological marking on the definite article allows gender to be assigned unambiguously. Therefore, to say that assigning gender from an article forces stereotype information to be considered is to say that this process activates information that can only complicate its own operation, as a clash with the stereotypical gender must be resolved in favour of the morphological gender. The Spanish data show that this clash *does* arise and *does* have to be resolved. However, it is more plausible to suggest that the stereotype information is accessed by processes that use information from nouns to construct representations of referents, including information about their (probable) gender, in mental models. Once this information is activated, other processes that assign gender cannot avoid encountering it, even if they do not "wilfully" activate it themselves.

A related issue is whether morphological gender marking might be processed more quickly by Spanish readers than by English readers, because gender marking is more important in Spanish than in English. The results of Experiment 4 show that in Spanish a clash between the morphological gender of a definite article and the stereotypical gender associated with a role name has an effect while the noun phrase is being processed. In contrast, evidence from studies of the interpretation of gender-marked pronouns in English (Gernsbacher, 1989; MacDonald and MacWhinney, 1990) suggests that the effect of this information on the suppression of the names of potential, but not actual, antecedents for the pronouns may be delayed. Whether English and Spanish differ in this respect remains to be definitively established. The probe name task used in these experiments is different from the self-paced reading task we used, and the particular case of morphological marking is different. In any case, the effect of gender information in these probe name studies shows up by the end of the clause in which the information occurs. Even if Spanish readers use morphological gender marking slightly more quickly than do English readers, that fact would not change the conclusions we draw in this paper.

There is one other explanation of our results that we must discount. It might be argued that our Spanish results could be explained by the printed frequency of the two forms of the role names. The Spanish nouns themselves usually, though not always, have different forms (e.g. *carpintero/carpintera*). The "mismatch" form (e.g. *carpintera*) will almost certainly be less frequent than the "match" form (e.g. *carpintero*), which could contribute to the slower reading of the sentence containing the mismatch form in the first sentence. However, if this were the explanation, we would expect a match/mismatch effect later in the passage, because the frequency effect does not involve stereotype information, and we

know from the English data that stereotype effects will show up late, if that is the first point at which stereotype information is used. But no late effects are found in Spanish, so the effect on the first sentence is unlikely to be explained by word frequency. In addition, the size of the mismatch effect in Experiments 2 and 3 is too large to be explained by frequency differences (Experiment 2, 235 msec; Experiment 3, 225 msec). However, the crucial effect in the first frame of the sentences in Experiment 4 is smaller (87 msec), and frequency effects could have made a significant contribution to it. To check for such effects, we divided the materials into those containing nouns with different forms and those containing nouns with the same masculine and feminine form, for which frequency effects cannot arise. The mismatch effect was 96 msec for the 8 items with identical masculine and feminine forms, $F(1, 7) = 16.46$, $p < .005$, and 67 msec for the 16 items with different masculine forms, $F(1, 15) = 10.85$, $p < .005$. As the effect was numerically larger and statistically more reliable for the items with identical form, we conclude that the mismatch effect cannot be explained by frequency.

Stereotypes can be the starting point for representations of particular characters in a text. If they are the only source of information about gender, as they are in the first sentences of our English texts, they can be used to encode elaborative inferences about gender into mental models. These inferences may, however, be overridden by later information. The fact that readers had difficulty comprehending the critical sentences in the mismatch conditions supports the claim that they attempt to construct a single coherent mental model around the main character (e.g. Garrod & Sanford, 1988; 1990). Information about main characters is foregrounded and updated by incoming information. Furthermore, although the assignment of gender in Spanish does not require a knowledge-based inference, because the gender is determined by the definite article, Spanish readers may encode a different inference into their mental models: that the gender of the character clashes with that of the stereotype.

Our results, therefore, appear to contradict the minimalist hypothesis (McKoon & Ratcliff, 1992). If an inference is made, minimalists must claim that (1) the subjects adopted a special strategy, or (2) the inference was necessary for local cohesion, or (3) the inference was based on easily available information. However, first, there is no reason to think subjects adopted a special strategy in these experiments, except to save the minimalist hypothesis. Second, if an inference (either about probable gender in English, or about a clash with the stereotype in Spanish) is made when the first sentences of our passages are read, it is an elaborative inference that is not necessary for local cohesion. Third, because there is no a priori definition of easily available information, minimalists could argue that information about stereotypes is easily available. But this argument makes the minimalist position circular and unfalsifiable, unless there is an independent criterion for what information is readily available in the required sense. Furthermore, although it is plausible to suggest that information about stereotypes is readily available, it is less plausible in the context of McKoon and Ratcliff's claims about other types of default information. For example, to accommodate experimental findings on instrument inferences, McKoon and Ratcliff are committed to the view that default instruments are not readily available. We are not claiming that information about instruments and information about stereotypes are the same. However, one must, at least, be

cautious about claiming that information about social stereotypes is readily available, if information about typical instruments (e.g. spoon used to stir coffee) is not.

To look at this matter from a different perspective, if we had not found differences between the match and the mismatch conditions, and if our experiments had sufficient power to detect such differences, we might have argued that our results supported the minimalist position. We could have said that stereotypic information is not used to make an inference about gender because such an inference is not necessary for local cohesion and because we had no reason to suppose that such information was available, in the required sense.

In relating our results to the minimalist hypothesis, it is important that the assignment of gender to a particular character in a text from a stereotype is an inference based on knowledge about the world—about what sort of people typically have which jobs. Otherwise the mismatch could be just a mismatch between linguistic features, as in the morphosyntactic mismatch condition in Experiment 2. Thus, for example, there is little doubt that readers would have difficulty in linking a feminine pronoun "she" to a character previously introduced by the proper name "John" (just as they had difficulty linking masculine "*el*" to feminine "*la enfermera*" in Experiment 2). However, in our culture, the link between proper names and the intended gender of their recipients is conventional. Thus, parents consider girls' names for their new born daughters and boys' names for their sons. Whether a name is for females or for males is regarded as a property of the name itself. So, although English proper names are for the most part not morphologically marked, they follow conventions that can conveniently be described as linguistic. However, in this sense it is clearly *not* a *linguistic* convention that "engineer" is a term applied to men. In addition, many Spanish names have morphological indications of the gender of their intended recipients (e.g. many female names end in *-a*, whereas -o and *-el* are common endings for male names). For a woman to be called "John" or "Pablo" is very different from a woman deciding to become an engineer.

We have argued that mental models of a text incorporate information about gender supplied by world knowledge, not only by the text, and that information may be updated by later information conveyed by, for example, the morphological form of a pronoun. Although spatial aspects of mental models have often been emphasized, such models are clearly more than just spatial representations. In addition to location information (e.g. Bower & Morrow, 1990; Morrow, Bower & Greenspan, 1989; O'Brien & Albrecht, 1992; de Vega, 1995), and spatial information associated with the main character (e.g. Glenberg et al., 1987), mental models also include information about other aspects of people, such as their goals (e.g. Huitema, Dopkins, Klin, & Myers, 1993), their emotional states (e.g. Gernsbacher, Goldsmith, & Robertson, 1992; Gernsbacher & Robertson, 1992), and their physical characteristics, attitudes, and beliefs (Albrecht & O'Brien, 1993). Our findings further show that readers have access to the characteristics of, or to a general profile of, the main characters described in a text.

REFERENCES

Albrecht, J.E., & O'Brien, E.J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *19*, 1061–1070.

Bower, G., & Morrow, D. (1990). Mental models in narrative comprehension. *Science*, 44–88.

Carreiras, M., Garnham, A., & Oakhill, J. (1993). The use of superficial and meaning-based representations in interpreting pronouns: Evidence from Spanish. *European Journal of Cognitive Psychology*, *5*, 93–116.

Chang, F. (1980). Active memory processes in visual sentence comprehension: Clause effects and pronominal reference. *Memory and Cognition*, *8*, 58–64.

Clark, H.H. (1977). Bridging. In P.N. Johnson-Laird & P.C. Wason (Eds.), *Thinking: Readings in cognitive science*. Cambridge, U.K.: Cambridge University Press.

Garnham, A. (1987). *Mental models as representations of discourse and text*. Chichester, UK: Ellis Horwood.

Garnham, A., & Oakhill, J. (1985). On line resolution of anaphoric pronouns: Effects of inference making and verb semantics. *British Journal of Psychology*, *76*, 385–393.

Garnham, A., Oakhill, J., Ehrlich, M.F., & Carreiras, M. (1995). Representations and processes in the interpretation of pronouns: New evidence from Spanish and French. *Journal of Memory and Language*, *34*, 41–62.

Garrod, S., & Sanford, A.J. (1988). Thematic subjecthood and cognitive constraints on discourse structure. *Journal of Pragmatics*, *12*, 519–534.

Garrod, S., & Sanford, A.J. (1990). Referential processes during reading: Focusing on roles of individuals. In D. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gernsbacher, M.A. (1989). Mechanisms that improve referential access. *Cognition*, *32*, 99–156.

Gernsbacher, M.A., Goldsmith, H.H., & Robertson, R.R.W. (1992). Do readers mentally represent characters' emotional states? *Cognition and Emotion*, *6*, 89–111.

Gernsbacher, M.A., & Robertson, R.R.W. (1992). Knowledge activation versus sentence mapping when representing fictional characters' emotional states. *Language and Cognitive Processes*, *7*, 353–371.

Glenberg, A.M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language*, *26*, 69–83.

Huitema, J.S., Dopkins, S., Klin, C.M., & Myers, J.L. (1993). Connecting goals and actions during reading. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 1053–1060.

Johnson-Laird, P.N. (1983). *Mental models*. Cambridge, UK: Cambridge University Press.

MacDonald, M.C., & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. *Journal of Memory and Language*, *29*, 469–492.

McKoon, G., & Ratcliff, R. (1992). Inference during reading. *Psychological Review*, *99*, 440–466.

Morrow, D.G., Bower, G.H., & Greenspan, S.E. (1989). Updating situation models during narrative comprehension. *Journal of Memory and Language*, *24*, 304–319.

O'Brien, E.J., & Albrecht, J.E. (1992). Comprehension strategies in the development of a mental model. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *18*, 777–784.

Sanford, A.J. (1985). *Cognition and cognitive psychology*. London: Weidenfeld and Nicolson.

Sanford, A.J. (1987). *The mind of man: Models of human understanding*. Brighton, Sussex, U.K.: Harvester Press.

Sanford, A.J., & Garrod, S. (1981). *Understanding written language*. New York: Wiley.

van Dijk, T., & Kintsch, W. (1983). *Strategies of discourse comprehension*. San Diego, CA: Academic Press.

de Vega, M. (1995). Backward updating of mental models during continuous reading of narratives. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *21*, 373–385.

Original manuscript received 6 February 1995
Accepted revision received 29 June 1995

APPENDIX 1

(British) English Role Names Used in Experiment 1

Mean ratings on 11-point scale: 1 = strongly male, 11 = strongly female

Male stereotyped role names		*Neutral role names*		*Female stereotyped role names*	
paratrooper	1.24	set designer	5.30	nurse	9.07
bricklayer	1.43	art historian	5.33	cleaner	9.17
footballer	1.57	artist	5.43	switchboard operator	9.27
scrap metal dealer	1.63	musician	5.43	baby sitter	9.33
butcher	1.77	paediatrician	5.43	canteen assistant	9.43
chauffeur	1.83	newsreader	5.52	fortune teller	9.47
firefighter	1.83	student	5.60	infant teacher	9.57
plumber	1.83	psychologist	5.73	receptionist	9.57
lorry driver	1.87	astrologer	5.80	childminder	9.63
soldier	1.90	swimmer	5.83	dental assistant	9.63
plasterer	1.93	fashion designer	5.93	typist	9.63
undertaker	1.97	novelist	6.03	secretary	9.70
porter	2.00	singer	6.13	beautician	9.77
judge	2.02	opera singer	6.43	prostitute	9.83
pilot	2.07	interpreter	6.53	nursery nurse	9.87
sheep shearer	2.17	bank clerk	6.57	housekeeper	9.97
electrician	2.20	traffic warden	6.57	midwife	10.27
farmer	2.27	welfare officer	6.68	au pair	10.53
carpenter	2.47	interior designer	6.70	nanny	10.57
taxi driver	2.50	physiotherapist	6.76	cheerleader	10.77

APPENDIX 2

Spanish Role Names Used in Experiments 2, 3, and 4

Ratings on 11-point scale: 0 = strongly female, 10 = strongly male

Male stereotyped role names

sepulturero/sepulturera	grave-digger	9.68	*zapatero/zapatera*	shoemaker	8.94
futbolista	footballer	9.37	*obrero/obrera*	worker	8.82
farero/farera	lighthouse keeper	9.37	*paracaidista*	parachutist	8.69
chatarrero/chatarrera	scrap metal dealer	9.37	*pescador/pescadera*	fishermen/fisherwoman	8.57
carpintero/carpintera	carpenter	9.31	*taxista*	taxi driver	8.43
chapista	panel beater	9.30	*presidente/presidenta*	president	8.02

Neutral role names

misionero/misionera	missionary	5.44	*cocinero/cocinera*	cook	4.69
escritor/escritora	writer	5.43	*intérprete*	interpreter	4.56
periodista	journalist	5.00	*doctor/doctora*	doctor	4.44
trapecista	trapeze artist	5.00	*locutor/locutora*	newsreader	4.44
artista	artist	5.00	*cantante*	singer	4.34
abogado/abogada	lawyer	4.93	*recepcionista*	receptionist	4.12

Female stereotype role names

enfermera / enfermero	nurse	3.12	*puericultora / puericultor*	paediatrician	2.18
mecanógrafa / mecanógrafo	typist	2.81	*telefonista*	telephonist	1.81
modelo	model	2.68	*comadrona / comadrón*	midwife	1.69
secretaria / secretario	secretary	2.62	*esteticista*	beautician	1.62
sirvienta / sirviente	servant	2.56	*tejedora / tejedor*	weaver	1.57
florista	florist	2.37	*bordadora / bordador*	embroiderer	0.57