A review and synthesis of research on comprehension of the masculine as a generic form in English

Nancy M. Henley
University of California

Department of Psychology
University of California
1282 Franz Hall, Box 951563
Los Angeles, CA 90095-1563
U.S.A.
nhenley@ucla.edu

Joselito Abueg

LiteracyPro Systems, Inc.
333 West Maude Avenue, Suite 108
Sunnyvale, CA 94085-4367
U.S.A.
jabueg@literacypro.com

Abstract
In this article we survey studies providing empirical tests of questions related to the comprehension of masculine generics; these are categorized by experimental paradigm. Findings are aggregated and weighted Stouffer combined tests are performed to assess the probabilities of the findings across studies. The evidence indicates that not only are masculine forms more likely to lead to male-oriented than to other responses, but also that this tendency is strongly evidenced in both children and adults, and that females are more likely than males to interpret masculine forms generically (p <.00001 for each combined test). The evidence is mixed on whether sex-neutral or sex-inclusive forms provoke more generic interpretation, but may favor the inclusive forms. Many children, and 10% or more of adults tested, were unable to explain the masculine generic rule satisfactorily. Processing-time studies indicate that a male-related response is more cognitively available than other responses following a masculine generic antecedent. Finally, there is a residual tendency to make male-oriented responses regardless of the stimulus.
Key words: masculine generic, sexist language, pronoun resolution.

Resumo
Neste artigo examinamos estudios que proporcionan tests empíricos de preguntas relacionadas coa comprensión dos masculinos xenéricos categorizados a través do paradigma

experimental. Ofrécese o total dos resultados e aplícanse tests combinados de Stouffer para avaliar as probabilidades dos resultados a través dos estudios. A evidencia mostra que non só as formas masculinas son máis propensas ca outras formas a conducir ó caso masculino, senón que esta tendencia queda evidenciada notablemente en nenos e adultos, e que as mulleres son máis susceptibles cós homes de interpretar as formas masculinas como xenéricas (p <.00001 para cada test combinado). Os resultados son mixtos no que respecta a se as formas con sexo neutral ou as formas con sexo incluínte provocan máis interpretacións xenéricas, aínda que se favorecen as formas incluíntes. Moitos nenos, e o 10% ou máis dos adultos do test, non foron capaces de explicar a regra do masculino xenérico satisfactoriamente. Estudios acerca do tempo de procesamento indican que unha resposta cara ó caso masculino é máis fácil cognitivamente ca outras respostas seguindo un antecedente masculino xenérico. Por último, existe unha tendencia residual a empregar respostas co caso masculino independentemente do estímulo.
Palabras clave: masculino xenérico, linguaxe sexista, elección pronominal.

1. Introduction

The question of masculine generics' ability to function generically, that is, to be understood as referring equally to females and males, has been a matter of controversy. By masculine generic[1] we mean the use of a masculine pronoun, noun, or suffix as generic, to refer to either or both female and male referents (e.g., *fireman*, *mankind, the average person*, *he*) and not males alone. This question and ones that stem from it are the focus of this article.

The masculine generic (MG) is one of the many forms of sex bias[2] in language, which may be broadly described as exclusion of women, stereotyping of both women and men, and deprecation of women (Henley, Hamilton & Thorne, 1984). Sex bias is in turn one of the subfields of the broad field of language, sex, and gender, which encompasses such concerns as literary analysis of how gender enters into written texts, linguistic analysis of the way women and men talk, sociolinguistic studies of how gender influences and is constructed through conversational interaction, psycholinguistic studies of how language users come to produce, understand, and respond to sex-biased language, and many others[3].

Where scholars of this field once divided it into two broad forms, sex difference and sex bias (Lakoff, 1975; Thorne & Henley, 1975), they now recognize the mutual

[1] We prefer to use the term *masculine generic* rather than *generic masculine* because in the latter, *generic* used as an adjective implies that the masculine form does work generically, which is subject to empirical test. *Generic* in the former term is used as a noun to denote a word used to reference a class of things, in this case human beings. There may be masculine, feminine, or unbiased generics.

[2] We use *sex* in this article to refer to social gender, to avoid confusion in the use of the word *gender*, which has linguistic as well as social/psychological meaning.

[3] Readers interested in the broad field of language and sex/gender can find other topics in and approaches to the field in such works as those of Lozano (1995), Cameron (1998), Pauwels (1998), Bucholtz, Liang & Sutton (1999), Romaine (1999), and Livia (2001).

influence and connectedness between these forms, seeing them as parts of the same phenomenon (e.g., Cameron, 1998). The close relationship of these topics is visible in the review we present here of a single aspect of sex bias. Despite the very specific delineation of the topic as one about how women and men are represented and perceived, the differing language patterns and perceptions of females and males assert themselves repeatedly.

There has been extensive research on the exclusivity of the MG, yet exactly what conclusions one may draw from this research has been unclear. The studies of inclusiveness had varied methods and findings and focused on different age groups and different research problems. Many people interested in this field would probably say that they know the MG doesn't work as a generic, but they may not know the answers to such questions as: if the MG biases thought toward the masculine, how do we know it's so? To what extent does this bias influence thought? Is it found by all research methods tried, or only some? Does it hold across the lifespan? Is it true for both reading and hearing the MG? Are men and women affected equally by the MG, or is it possible that women are exempt from its influence? Can using alternative pronouns (e.g., *he or she*, *they*) lead to greater inclusion of females, and if so, which type works best?

There has also been much research on related topics, such as the differences between females' and males' understanding and production of the masculine generic, children's understanding and use of it, effects of and on cognitive, social, and personality processes, practical consequences of its use, and the utility of alternative terms and constructions. This research has been carried out in and on a variety of languages with different types and extents of grammatical gender (e.g., Gathercole, 1989; Jaworski, 1989; Rothermund, 1998; Heise, 2000; Doleschal & Schmid, 2001; see in particular Hellinger & Bussmann, 2001, which has chapters on gender in nine different languages).

We should note that the use of generic forms involves both prescriptive and descriptive grammar: the masculine is in many languages, including English, the grammatically prescribed form for generic usage. If it is not interpreted generically by readers and hearers, that is a fact of language that would make up part of the description of the language. To interpret the masculine generic (MG) nongenerically is not "incorrect", linguistically speaking, except in the prescriptive sense.

Some observers may say that the MG is hardly used any more, and is no longer worth pursuing. It's true that there has been dramatic change not only on the individual level, but in institutional use of the MG, as numerous arbiters of language, including style manuals, news providers, teachers, professional associations, textbooks, government publications, and popular media, moved to eliminate the most obvious forms of sex bias in their output and institutional culture (Cooper, 1984; Ehrlich & King, 1992). These changes are due in large part to feminist efforts at language reform, through such means as lobbying, letters to editors, interpersonal

interaction, teaching, research, and writing (see Pauwels, 1998 for thorough coverage of issues of language planning and change).

Despite that, and the fact that the singular use of *they* has long been common in vernacular speech (Bodine, 1975; Meyers, 1990), the MG is by no means out of use; it still appears regularly in written material, though its use may vary by source, and be especially resistant to change in some quarters. Cooper (1984) examined a corpus of 525,000 words from American publications from 1971 to 1979. During that time, although there was an overall decline in the use of the MG, it varied by publication; for example, there was no decline in prepared remarks, and little decline in extemporaneous remarks, by members of the U.S. Congress (as printed in the *Congressional Record*).

The focus of this article will, for reasons of length and manageability, be limited to empirical studies of how the MG is comprehended in English, and will exclude studies that do not directly address the issue of comprehension or compare different generic forms, and studies examining detrimental effects beyond comprehension[4]. We survey the literature from the first published study known to us, in 1971, through those of the 1990's; studies are presented in chronological order within a research topic. Much of this basic research was carried out in the 1980's. More recent research on this topic has developed in different directions, for example in examining pronoun usage within the context of analysis of literary texts (Livia, 2001), or investigating cognitive processing of the MG as part of the larger enterprise of understanding language processing in general (see Friederici & Jacobsen, 1999; and Schriefers & Jescheniak, 1999, for reviews).

In the hope of organizing and clarifying a varied and somewhat piecemeal literature, we will present a systematic review of the MG comprehension research, with meta-analytic tests to synthesize findings for each of several questions, and some discussion of methodological and theoretical issues.

2. Studies testing the comprehension of the MG

This field of research has a central problem at its core: researchers want to know what people think when they read or hear the MG, but none can know for sure what another person is thinking. If you ask the person directly, they could be unsure, or misrepresent what they were thinking, through either misunderstanding or intentional misdirection (for various reasons). Therefore researchers must infer thought from observations of the reader/hearer's external responses, and they have

[4] Some studies may be considered by their authors to be tests of the inclusiveness of MG forms (e.g., Cole, Hill & Dayley, 1983), but were not by us (see Henley, 1989: 65, for elaboration on this point; also see Gastil, 1990: 631-33, for further critique).

tried to do this in myriad ways –often indirect ones so that the participants won't be motivated to respond in politically correct or experimenter-pleasing ways.

Because the studies to be reviewed represent a rather wide range of methods, and different critical comments apply to different methods, they will be described here individually, though briefly, and organized by experimental paradigm. All results reported are significant with a probability of = .05 unless otherwise stated.

In reviewing the research, it will be useful to know the distinction between two types of sex-unbiased alternatives: *sex-neutral*, that is, either not marking sex (e.g., *they*, *one*, *people*) or not using a potentially sex-marked form at all (as in "The careful writer checks the facts", rather than "his facts"); and *sex-inclusive*, that is, marking or naming both sexes (e.g., *he or she*, *men and women*). These will be referred to collectively as *unbiased* forms.

2.1 Studies requiring identification of an antecedent's sex

The first research paradigm presents stimuli of some sort and asks research participants to tell the sex of the person featured in the stimulus. In the earliest empirical study found, Kidd (1971) presented MG sentences to college students and asked them to give, among other things, the antecedent's sex. Out of 1,017 responses, 92% were "male", and only 8% were "female". Although this finding suggests that a bias in interpretation exists, this study had no control condition in which responses to unbiased generic sentences were obtained. Maybe people would respond the same way in the absence of the masculine form, simply because they are more used to thinking of males, in a male-biased culture. The remainder of the studies to be reported do have comparison groups.

MacKay's (1980) research is different from the other research reviewed in this paper, because he gave comparison groups different made-up alternatives rather than existing alternative forms. Many English speakers over the years have called for a new, inclusive pronoun to replace the MG, and some have even created them (see Baron, 1981, 1986). This research responds to that call. MacKay had participants read two paragraphs which contained sex-neutral antecedents (e.g., *student*, *character*) with either the pronoun *he* or a neologism (*e*, *E*, or *tey*, given with no prior explanation), and asked whether the antecedent was male, female, or either. When *he* was used, the probability of including "female" in the response was .50. When the pronoun was a neologism, the probability of including "female" was .87. While the finding that a new and unknown word can provoke more inclusion of females than the MG is quite interesting, its usefulness is limited unless *e*, *E*, or *tey* is under consideration for adoption.

Some studies have used longer experimental materials –paragraphs or whole stories– rather than sentences, in order to come closer to participants' real-life experience of language, as well as give them longer exposure to the stimuli. Hamilton & Henley (1982; see also Hamilton, 1991, 1997) had participants listen to

an audiotaped story about a group of people, which was related using masculine (M), neutral (N), or inclusive (I) forms (each participant heard only one form). When asked later to respond to questions about the sex(es) of persons in the story, those who had heard the MG story were significantly less likely to answer that the stories were about both sexes than those who had heard one of the unbiased (U) generic stories.

Wilson & Ng (1988) used sentences as stimuli, but introduced a new method of inferring thoughts relating to the MG: they paired M or feminine (F) generic sentences with photographs of faces presented to participants at speeds below their previously-established threshold for discerning sex, and asked participants to state the sex of the persons in the photos[5]. Since the participants could not in fact make out the sex, the researchers assumed that if pronoun gender causes thoughts of one or the other sex, it would affect the identification of photos. They found that significantly more male face responses were reported in the MG than in the F condition.

Wilson & Ng make the interesting suggestion that one could use the phenomenon of binocular rivalry resolution to study the MG, presumably with a similarly-structured experiment but, following a generic stimulus, presenting a female and male picture to the different eyes simultaneously at fast speeds. Participants would then report the sex of the person that they thought they saw. Wilson & Ng imply that this method would get closer to ascertaining the viewer's imagery than methods already in use, which only allow the experimenter to infer imagery from reports after the fact.

One of the authors of the previous study, Ng, in another series of studies (1988, 1990), used a subtle way of having participants indicate how they had interpreted the MG, by using the phenomenon of proactive inhibition. This is the tendency for memory to deteriorate across trials when people are repeatedly given words from a single category (e.g., "fruits") to recall. However, switching to a new category leads to release from proactive inhibition, that is, restored ability to recall the words. In his Experiment 1, Ng gave a set of either feminine or masculine word pairs –a generic class and a person's name, such as "girl/Iris", "son/Lewis"– to his participants (each participant had only one type of word pair, i.e., feminine or masculine), and after a brief distracter task asked them to recall (aloud) the stimuli. As expected, their recall declined over the four trials of similar word pairs. Both groups were next given the same stimuli: MG words paired with sex-neutral names (man/Robin, his/Chris). Recall continued to decline for those originally in the M condition (continuation of inhibition), as would be expected if they had interpreted these MG word/N name pairs as more M stimuli, similar to the ones they had previously had. But for those

[5] The task may thus be construed as one of assignment of sex to an antecedent rather than of actual identification of the person's sex.

originally in the F condition, recall improved (release from inhibition), as would be expected if they too had interpreted these MG word/N name pairs as M stimuli, dissimilar to the F ones they had had. In other words, both groups interpreted the "masculine generics" as referring to males, not neutral, and applied this interpretation to the sex-neutral names that followed.

In summary, the studies in this section requiring sex identification following presentation of the MG, all showed a bias towards the masculine.

2.2. Studies requiring generative response

Several studies have required participants to generate verbal or graphic responses identifying sex of referents –to draw pictures, to assign names, or to write stories after presentation of stimuli.

Many studies have used students as research subjects, not simply because they are available, but because the researchers are concerned to find out what influence the MG may have on those they are trying to educate. Harrison (1975) is one. She gave junior high school students a survey asking their knowledge of early humans, in a form which used only one way of referring to human beings, in M, N, or I terms. Following this, she asked the students to draw pictures of the early humans. Those participants given the M form included females in only 26% of their drawings, but those given the N or I form included females in 90% and 98.5%, respectively.

Working with slightly older subjects, Shimanoff (1975, cited in Shimanoff, 1977) required written responses. She presented generic stimulus sentences in MG or one of two N generic forms, and had participants write paragraphs about them. The MG elicited significantly more writing about a male than did either of the N forms (actual percentages are not given). Similarly, Moulton, Robinson & Elias (1978) asked college students to write a story based on a theme presented with M, N, or I pronouns. With the M pronoun only 35% of the story characters were female, but with the N and I forms the percentage of female characters increased to 46-56% (respectively).

The first extensive series of experiments in MG comprehension and production was carried out by Martyna (1978; see also Martyna, 1980b), using a variety of experimental paradigms. In her Experiment 5, Martyna asked college students to generate names for the subjects of sex-indefinite generic sentences which were presented with one of the M, N, or I pronouns, for example, "someone who always turns in [*his*/*their*/*his or her*] homework late". With the M pronoun, 88% of the names generated were masculine names, while with the U forms only 62-67% were.

Researchers using younger participants, not as proficient at reading and writing as adults, have often had them draw, point to, or speak, rather than write, their responses. Hyde's (1984, Experiment 1) subjects had a wide range of ages, and she tailored her research methods to their ages. Hyde presented stimulus sentences to elementary and college students using M, N, or I generic pronouns, and asked them

to tell (in the case of elementary school students) or write (in the case of college students) a story based on the stimulus sentence. Stories in general tended to be about males, but significantly more stories were written including female characters when the U forms were used (*he or she*, 42%; *they*, 18%) than when the M form was (12%). In a similar experiment with children only (Experiment 2), Hyde gave participants stimulus sentences with the same three pronoun conditions and with *she* as a pronoun, as well. Female characters produced in their stories were 31% with *they*, 18% with *he or she*, 17% with M forms, and 77% with F ones.

Another way to try to discern the subtle influence of pronoun gender on responses is to have other persons judge what participants produce (without knowing the conditions the participants were in), and not only to identify sex but to determine the degree of masculinity/femininity of the response. Sniezek & Jazwinski (1986, Experiment 1) did this. They asked college students to draw a member of the group labeled (according to experimental condition) *man*, *people*, or *men and women*; independent judges then rated the pictures on a 7-point scale from most feminine to most masculine. The pictures stimulated by the M noun were rated significantly more masculine than those stimulated by the N one; ratings of pictures stimulated by the I phrase fell between the two.

An even more indirect approach was used by the same researchers in another aspect of this research: comparing actuarial estimates made by participants assigned to different pronoun conditions. These researchers had participants read a brief essay on the longevity of *man*, *human beings*, or *men and women*, in which other nouns and pronouns were in keeping with the generic noun condition (each participant in only one condition). After this, participants were to estimate human life span, and later, to estimate life expectancy for females and males separately. The researchers assumed that the closer the generic life expectancy estimate was to that for a specific sex, the more the participant's generic form stimulated thought biased toward that sex. As expected, with the MG, estimates for the generic group longevity were closer to those for males than for females, while with N and I the generic group estimate was less biased, and toward the female side; however, the condition effect was only marginally significant ($p<.06$). Since the connections between exposure to the pronoun and estimates of life expectancy were rather tenuous (and possibly influenced by prior knowledge), it is not surprising that the difference found did not reach the standard level of significance; that they came so close in this long-shot research is surprising in itself.

Like Martyna (1978), Sniezek & Jazwinski (1986, Experiment 1) had college students provide names, in this case for people in 15 occupations. Each occupation was given a brief definition written in M, N, or I language, with each participant in only one pronoun condition. Names generated by participants in the M pronoun condition were 93% masculine, a significantly greater proportion than the 76-77% masculine names generated by participants in the I and N conditions.

Using the now-familiar method of sentence fragment completion, Hamilton (1988), in an experiment described in more detail in the section on imagery below, had college students complete sentence fragments using M or U pronouns without explicitly telling them to. They were then asked to generate names for the persons their sentences were written about. Those producing the M pronouns gave relatively more masculine names (ratio of M:F names = 3.5:1) than those producing U pronouns (M:F names = 2.0:1); this pronoun effect, however, was only marginally significant ($p<.10$).

Khosroshahi (1989), like Harrison (1975) and Sniezek & Jazwinski (1986, Experiment 1), looked at the sex of figures participants drew in response to paragraphs that were written in M, I, and N styles (each was given all three types). Participants themselves identified the figures' sex. For the I paragraphs, participants produced significantly more female (34%) and fewer male figures (54%) than they did for the MG ones (female 19%, male 67%). The sex of figures drawn for N paragraphs did not differ significantly from that of those drawn by participants for either of the other two pronoun types.

Switzer (1990), like Hyde (1984), asked young participants to complete a story. She directed first-grade and seventh-grade children to write an ending to a story about a generic student's first day at school, given the beginning of the story worded in M, N, or I pronoun, or noun repetition (no pronoun) form. She found significant differences due to stimulus condition, with 93% of students telling stories about males when given the M pronoun, but only 44-49% doing so when given the N or I pronoun; 63% described a male when in the noun repetition condition.

It's worth noting that the Switzer study, and Hyde's as well, used stimuli that were not necessarily generic. In Hyde's Experiments 1 and 2, the interviewer's preamble to presenting the stimulus to elementary school children is, "I'm going to tell you a little bit about *a kid*, and then I want you to make up a story about *that child*" (Hyde, 1984: 699; emphasis added). Thus, the stimulus sentence "When a kid goes to school, [*he*/*they*/*he or she*] often feels excited on the first day", when paired with *he* or *she*, may well be taken specifically rather than generically. In Switzer's experiment, the young participants were told to "pretend that *a new student* is coming to be a part of your class. Tomorrow will be [*his*/*his or her*/*their*/*the student's*] first day. Describe how you think [*he*/*he or she*/*they*/*the student*] will feel –you can make up a name for [*him*/*him* or *her*/*them*/*the student*]" (Switzer, 1990: 74; emphasis added). The singular antecedent here too may be taken specifically rather than generically when paired with the singular pronoun. One could argue that that is just the point: gendered pronouns are understood specifically, not generically. However, that argument is shaky if the rest of the sentence is open to nongeneric interpretation. Hence the results from these two studies need to be interpreted with some caution.

To avoid problems like these, MacKay & Fulkerson (1979) gave excellent guidelines for constructing truly generic sentences: their sentences combined a

generic pronoun with a habitual action marker (e.g., *usually*), were written in present tense, and had an indefinite antecedent noun preceded by the article *a* (e.g., "A nurse must frequently help his patients get out of bed").

In the main, results with generative responses, like those with direct identification of stimulus sex, showed a general influence of the MG to skew productions toward the masculine. Only Sniezek & Jazwinski's (1986) actuarial estimates and Hamilton's (1988) results with naming did not reach statistical significance, though other methods used in the same studies (for Hamilton, see section 2.5) did show significant influence of the MG.

2.3 Studies requiring a response to be chosen from an array

Studies requiring a choice from an array have put both female and male pictures or names (usually in equal numbers) before the participants, which their own mental generating capacities may not do. This reminding function might be expected to counteract a cultural predilection to think of humans as males, and stimulate equal female and male responses, if a word truly functions generically, or at least to provoke more responses designating females than in the generative-response studies.

In an early study in this category, Schneider & Hacker (1973), gave college students in two different sections of the same sociology course different labels to represent the major divisions of the text, and asked them to bring in pictures clipped from the print media to represent these divisions (we classify this as selection from an array, since the students had to choose from the finite population of printed pictures.). With M wording (e.g., "Social man") 64% of the pictures showed only males; with N wording (e.g., "Social behavior") only 50% of the pictures submitted showed only males. However, the high percentages of male-only pictures may have been influenced by the higher incidence of such pictures (relative to female pictures) available in print, since the "array" in this case was not controlled. Later research used controlled arrays.

Experiments with young people have typically used pictures for the arrays. Harrison & Passero (1975) gave third-graders eight statements about human activities in either M or N generic form, and asked them to choose an illustrative picture from an array. M-worded statements elicited exclusively male-oriented responses from 49-85% of the students, while the N-worded ones elicited only 3-31% all-male responses (percentages are the range for individual stimulus statements; aggregates were not reported). Eberhart (1976/1977) had children respond to statements which used either M or N generic nouns, by indicating pictures which would illustrate the statements. The pupils indicated pictures including females significantly more often (97-98%) when the N forms were used than when the M ones were (18-29.5%).

Working with college students, Martyna (1978) in her Experiment 6 followed the same pattern as her Experiment 5 (described above), except that rather than

generating names, participants chose names for subjects of sentences from an array. Sixty-one per cent still chose masculine names with the M pronoun, while the U forms led to fairly equal choice of masculine and feminine names (51-52% masculine), a significant difference.

DeStefano, Kuhner & Pepinsky (1978; see also DeStefano, 1978/1979) used picture arrays with both high school and college students. The researchers presented sentences containing M or N terms for people, and had participants indicate the referent of the sentence by choosing an illustration from an array (which included both singular and plural single-sex and mixed-sex figures, and split ones to indicate indefinite sex). With the generic pronoun *man* or *men*, most referent choices were singular or plural male figures, while for *human being* or *people*, mixed-sex illustrations were most often chosen.

Other investigators have used picture arrays with adults also. Wilson (1978) tested elementary and secondary school teachers for their inclusion of females and males in understanding generic nouns given in either M or N form; they were asked to respond by indicating all pictures which applied to the sentences given. The pictures were of only males, only females, or both females and males. The M form elicited male-only responses significantly more often (96%) than it did both male and female (75.5%) or female-only (71%) responses; however, the N form elicited all three types of response at high rates (97%, 96%, and 98%, respectively).

In his series of experiments with masculine generics, Ng (1988, Experiment 2) gave young participants neutral generic stimulus sentences, and had them generate tag questions, which typically involved generating a pronoun. Then they were asked to select a picture from an array to illustrate the sentence (thus they were responding to their own stimuli, not ones given to them). Those who had made their tags with M pronouns chose a female picture only 33% of the time, much less often than those who used N (68%) or F pronouns (81%), or used some other form (62%).

The research in this section found that choosing from arrays, like identifying sex or giving a generative response, was influenced by pronoun gender even though participants had alternative interpretations presented to them in the arrays.

2.4. Studies requiring participants to indicate whether a given picture is appropriate

In two of her studies (3 and 4), Martyna (1978) presented generic-topic sentences using *he*, *they*, or *he and she* to participants at very fast speeds. They were then shown a picture of a female or male and required to make a fast Yes/No decision on whether the picture was appropriate to the sentence. When *he* was used, participants said no to the female picture, that is, that the MG could not be referring to a female, in 11 to 25% of the judgments (Experiments 3 and 4, respectively). The correct answer, if *he* were acting generically, is “yes”, so there should have been 0% “no” responses (responses to the other pronoun conditions are not reported).

MacKay & Fulkerson (1979, Experiments 1 and 4), in a paradigm similar to Martyna's, presented stimulus sentences using generic *he*, aurally rather than visually, and had participants judge whether the sentence could refer to one or more females, by pushing a "yes" or "no" button (here too all correct answers were "yes"). They were wrong, that is, incorrectly excluded females, 87% of the time. When, however, a generic noun (e.g., *student*) was used, they erred only 43% of the time.

These two studies both found that exposure to MG forms led to excluding females, though the exclusion rates differed considerably. This variation, visible in experiments in other sections as well, could be due to differences in experimental method, subject populations, or other factors.

2.5. Studies examining imagery evoked by generic stimuli

Some experiments have asked participants about the imagery they experienced when responding to different generic stimuli; however, some studies (e.g., Schneider & Hacker, 1973; Switzer, 1990) have merely inferred participants' imagery, without having asked directly about it. Here we cover only those studies which have explicitly asked about imagery. The reasoning behind these experiments is that if MG pronouns lead to more mental images of males, then the pronouns are biasing those persons exposed to them to think of males more.

As part of her series of experiments, Martyna (1978) also investigated imagery. In Experiments 1 and 2, child and college student participants completed sentence fragments which required a pronoun to be used, and were asked about any imagery they experienced when generating them (here again participants are responding to their own stimuli). As expected, when the subjects of the sentences were stereotypically related in content to a particular sex (e.g., *engineer*, *secretary*), high percentages of the generated pronouns (87-96%) and experienced imagery (90-92%) reported were of the associated sex. More relevant to our purposes are the responses to neutral-content fragments: 65% of neutral fragments were still completed with *he*, 30% used N or I alternatives, and 5% used *she*. The associated imagery reported with these neutral antecedents was 51% male, 39% neutral, and 10% female.

Note that Martyna's Experiments 1 and 2 examined the circumstances under which different pronouns were produced (which we do not cover in this review), more than the comprehension of pronouns; we include them only for the imagery reported with the pronouns. However, we should note that we can't conclude that the pronouns were responsible for the described imagery in this early research, since the gendered pronoun and the sex-stereotypical content were essentially confounded; these studies mainly showed an association among content, gender of pronoun, and reported imagery. Where higher male than female imagery was reported with higher production of M pronouns, it could have been due either to the male-related content of the sentence or, with neutral content, to an underlying masculine bias influencing both imagery and pronoun choice. It's also possible that the imagery, influenced by

content, preceded and influenced pronoun choice.

A different method was used by Schaeffer (1982), in a large-scale experiment which was part of the 1980 General Social Survey (conducted by the National Opinion Research Center). Respondents were asked to evaluate the general desirability of 13 traits for a child. Half the 971 respondents were given the trait descriptions in M pronoun form and half were given them in N noun form. After their response, the participants were asked to tell whether they had been thinking mostly about boys, girls, or both. Given the M wording, 74% reported thinking about both girls and boys, but with the N wording, 83% did.

In a similar paradigm, Hamilton & Henley (1982) had participants listen to a taped story using M, N, or I forms. When participants were asked afterward to describe any mental imagery they experienced, those in the N or I pronoun condition reported a significantly higher percentage including female images (ratio of male:female images = 1:.89) than did those in the M pronoun condition (ratio of male:female images = 1:.54).

Hamilton (1988), like Ng (1988, described in section 2.3), used non-sex-associated stimuli to elicit participants' responses to their own pronouns. However, Hamilton confronted the problem of individual differences that arises when participants generate pronouns freely: even with neutral stimuli, there is the possibility that some participants produce both the MG pronoun and masculine imagery because of a propensity to think of males, rather than producing the imagery as an effect of the pronoun. Experimenters like to use random assignment to conditions as a way to get around such personality differences affecting the phenomenon under observation. Hamilton's solution was to assign participants randomly to conditions of producing M or N forms, but without telling them to do so, by instructing them to complete generic sentence fragments using what she called either "traditional, formal" (typically MG) forms, or "modern, informal" (typically unbiased) forms. Those producing M pronouns reported significantly more male imagery (three times as many responses without females as with them), than participants producing the U forms (about even proportions including and excluding female images).

Hamilton also showed that this difference could not have been due to there being something more inclusive about the idea of "modern, informal" usage: in two other conditions of the experiment, the sentence completions were made with the same two sets of instructions, but with plural subjects in the sentences. Thus, pronouns had to be *they* in every case. Sex of names assigned, and imagery experienced, did not differ significantly in the plural "formal" vs."informal" conditions, thus the differences found with the singular conditions could be due only to the pronoun difference, not to the concepts of formality/informality, nor to individual propensities to think of males.

Describing imagery stimulated by experimental sentences was the method used by Gastil (1990). He had college students read a stimulus sentence aloud, then describe any image that came to mind. The sentences were written in generic form using *he*, *he*/*she* or *they* for pronouns (each participant was exposed to only one pronoun type). After all descriptions, participants were asked about the sex of persons in the images for each sentence. Significantly more images of males (for both female and male participants) were reported with the M than with the I or N pronouns, and significantly fewer female images with the M pronoun than with the I. More mixed images were also reported with the I and N pronouns than with the M; female and male participants showed this significant pronoun effect both separately and together.

The results found by the studies in this section showed that imagery reported, like responses described in other sections, is significantly biased toward the masculine by the use of the MG.

2.6. Processing time

If people primarily interpret MGs as referencing males, presumably they have males (or masculine things) in mind when they read or hear an MG, and it should take them longer to perceive a female-related thing than a male-related one. There were occasional early reports that it takes longer for readers to process the MG (generically) than it takes for neutral generics, or that people respond to male stimuli faster than to female stimuli following presentation of an MG.

One such was Martyna's (1978, Experiment 3). In it she measured reaction times in participants' judgments of whether a picture of a female or male was compatible with a sentence which contained an M, N, or I generic pronoun. Martyna found that after reading generic *he*, participants (especially male ones) responded to pictures of males faster than to those of females; but with the other pronouns, there was no significant difference in time to respond to female or male pictures, as would be expected if they are unbiased (though this finding was complicated by sex differences).

Silveira (1978) conducted a similar experiment, using generic M or N nouns (*man*, *people*), with female participants. She too found that it took longer to respond to female pictures when the antecedent was M rather than N. MacKay & Fulkerson (1979) also obtained reaction times in their experiments asking whether a sentence (generically-worded, using generic *he* or an alternative) could refer to one or more females. They found longer times to respond (especially correctly, that females were referred to) when *he* was used than when a generic noun was used.

Hamilton & Henley (1982) timed two types of their participants' responses to the story they heard related in M, N, or I form: time before hitting a button to indicate a light had flashed while they were listening to the story, and time to respond to each true/false question about whether males and females were included in the story (for either, a slower time to respond when the M was used would indicate longer processing time, presumably due to the added cognitive processing burden).

While they found that female participants took longer to respond to both lights and questions with the M than with U forms, and male participants took longer with U than with the M, the interactions were not significant.

These early studies of reaction time differences often were not a primary focus of the research, sometimes had equivocal results or small numbers of participants, and were carried out before computers became widely available to pace and time the research procedures. As more has become understood about cognitive processing, and more sophisticated equipment has been available for measuring reaction time and running such experiments, reaction time experiments have become somewhat more systematic. Some of these have built on previous findings of longer processing time when an antecedent and pronoun were mismatched (e.g., Carreiras et al., 1996), but they often did not utilize generic stimuli.

Banaji & Hardin (1996, Experiment 1) did have a relevant contrast between masculine-related and neutral generic terms. In these experiments generic nouns, some of which were sex-associated (e.g., *nurse*, *doctor*) and others not (*postal*, *clerk*), were used as primes. These were exposed on a computer screen so fast that they were below the threshold of awareness, then were followed by presentation of a target M or F pronoun (e.g., *his*, *she*). Participants were to judge the sex of the pronoun and press a key marked M or F as quickly as possible. Besides reporting a noun/pronoun gender-matching effect indicating automatic sex stereotyping (and speeds approximating the gender-matched speeds following neutral primes), the researchers separately analyzed MG primes, words containing the morpheme *–man*. After MGs, correct judgments were slower when the target pronoun was female than when it was male, indicating that the generic term stimulated a masculine interpretation. Even excluding terms that are sometimes used specifically (e.g., *fireman*), to analyze only those clearly meant generically (e.g., *mankind*), they found significantly faster correct responses for male than female pronouns.

Foertsch & Gernsbacher (1997) used another common psycholinguistic paradigm: they had participants read sentences on a computer, with the participants controlling the speed at which the clauses were presented. Each sentence had a stereotypically masculine or feminine, or a neutral or indefinite, antecedent used in a generic fashion, followed by a singular pronoun in the next clause (e.g., "Anybody who litters should be fined $50, even if [*he*/*she*/*they*] cannot see a trashcan nearby…"). Clauses paired with gender-mismatched pronouns were read most slowly, while those with same-gender pronouns or the singular *they* were read with equal facility. When the antecedent was neutral (e.g., "A runner"), the reading times for *he*, *she*, and *they* were equal; when the antecedent was indefinite (e.g., "Anybody"), *they* clauses were read faster than either *she* or *he* clauses.

These results seem somewhat mixed: the fact that generic *they* clauses were not processed as gender-mismatched, as other-gendered pronoun clauses were, and were read more quickly than gendered pronoun clauses, is evidence that *they* may be

interpreted generically; but there was not evidence that the M pronoun impeded processing any more than the F pronoun did. But then the F pronoun is not the one that is the prescriptive generic. The results give evidence that with a generic antecedent, singular *they* is interpreted more generically than the MG.

These findings of additional time needed to process the MG generically are further evidence that it is ordinarily processed as masculine (though some early findings were mixed or nonsignificant). Note that the findings also suggest that readers who process the MG as masculine might have swifter and less difficult reading than those who are interpreting the MG as it is supposed to be interpreted. On the other hand, when a correct response is dependent on correctly processing MGs, the MG unduly burdens the reader, and U generics ease the burden, allowing a swifter correct reading.

This burden of processing the MG may cause additional detrimental effects, such as in comprehension or recall of surrounding content, especially for those who do make the effort to include females: Crawford & English (1984) found poorer recall of the factual content of an essay among female college students, especially good learners, when it was written with M than when with I generic forms.

3. Overview and quantitative integration of inclusiveness studies

In summary, in the significant findings of every research paradigm, with both children and adults, males and females, the MG form was predominantly found to bias understanding toward the male. The percentages of female inclusion given the M form varied greatly between studies, even with similar paradigms; however, in no studies was female inclusion equal to male when the M form was used as a generic. This masculine bias is largely true also for the N and I forms; however, these alternatives to the MG provoked significantly more inclusive responses, in nearly every case, than the MG did.

The pervasiveness of masculine-biased response, even in neutral conditions, which we attribute to cultural bias, will be discussed later in this paper. This pervasive bias, however, does not mean there is no bias attributable separately to the use of masculine language forms. Since masculine cultural bias is present when participants are given U forms as well as M ones, the difference between the inclusiveness found with the M and that found with U forms is an indication of what masculine language adds to existing cultural sex bias. This difference is significant if the test for a main effect of form (M, N, I, etc.) on the inclusiveness of responses is significant.

3.1. Meta-analyses

The results of these studies were subjected to meta-analysis. Meta-analysis is a general term encompassing various quantitative methods for synthesizing research by combining information from different studies of the same phenomenon (Wolf,

1986; Rosenthal, 1991). In order to integrate the findings of different studies, a common metric was chosen which was applicable to the greatest number of studies: percentage of female-including responses made within each pronoun condition. Nineteen studies, in 17 different papers cited above, were both amenable to this metric and reported statistical tests; these are listed in Table 1[6].

For convenience, for the meta-analyses the studies will be referred to by the numbers in this table. Some of the studies, such as 7, provided the percentages directly; others, such as 11, provided percentages for female-excluding rather than female-including responses, and were easily convertible by subtraction from 100%. Still others, such as 1, 3, and 4, required more elaborate transformations to yield percentage of female-including responses: these were retrieved from percentages, ratios, etc., sometimes separately given for different groups of participants, in combination with the given numbers for the different groups. Figures 1, 2, and 3 show the comparisons of the M form with the N, I, and U forms included in the different papers, using this common metric[7].

Because reported measures and statistical tests among these studies varied widely, and means and standard deviations were rarely reported (hence effect size could not be calculated), the only meta-analytic test deemed applicable was the Stouffer combined test (Stouffer et al., 1949: 49; Wolf, 1986: 20; Rosenthal, 1991: 68-70), which requires only the *p* values of the studies under consideration. Since *N*'s were available for all of the 19 usable studies, a weighted version of the Stouffer combined test could be calculated, as suggested by Mosteller & Bush (1954) and Wolf (1986: 39-40). This test combines the probabilities of the various studies[8] and weights them by the number of participants in each, yielding a standard normal deviate, or *z*[9] value.

---

[6] Only one measure was used from each experiment included. DeStefano, Kuhner & Pepinsky (1978) and Shimanoff (1977) were not amenable to this metric; Kidd (1971) could not be used because there is no comparison of conditions; and Schneider & Hacker (1973), because no test of probability was reported. Martyna's (1978) Experiments 1 and 2 were not a clear test of the effect of pronoun type. The various processing time experiments had widely varying tests asking different questions (e.g., comparing different stimuli or different responses), or had problematic analyses (e.g., including specific with generic stimuli, or not analyzing for main effect of generic type), or omitted necessary information, so that we could not include them in the meta-analyses. Harrison & Passero (1975) could not be graphed because aggregate figures were not given (only responses to individual items).

[7] Although Schneider & Hacker (1973) could not be included in the meta-analysis because information was missing, the study did give percentages so that the results could be graphed; therefore they are shown in Figure 1.

[8] In cases in which the *p* value was not given but the result was said to be significant, *p* was conservatively set at .05. In cases where no statistical test of a particular finding was reported, *p* value was conservatively set at .5.

[9] The statistic *z* indicates how far distant from the mean an observed value of a distribution is, in standardized terms (the unit of measurement is the standard deviation of the distribution). Thus the larger the *z*, the farther away from the mean, and the more unlikely (hence "significant"), it is; a *z* of 1.64 has a probability of about .05.

Table 1. *Studies used in meta-analyses.*

Author/Year/Study

1.Eberhart (1976/1977)
2.Gastil (1990)
3.Hamilton (1988)
4.Hamilton & Henley (1982)
5.Harrison (1975)
6.Harrison & Passero (1975)
7.Hyde (1984, Expt. 1)
8.Hyde (1984, Expt. 2)
9.Khosroshahi (1989)
10.MacKay (1980)
11.MacKay & Fulkerson (1979, Expts. 1/4)
12.Martyna (1978, Study 5)
13.Martyna (1978, Study 6)
14.Moulton, Robinson & Elias (1978)
15.Ng (1988, Expt. 2)
16.Schaeffer (1982)
17.Sniezek & Jazwinski (1986, Expt. 1, names)
18. Switzer (1990)
19. Wilson (1978)

Figure 1. *Masculine vs. Neutral Comparison*

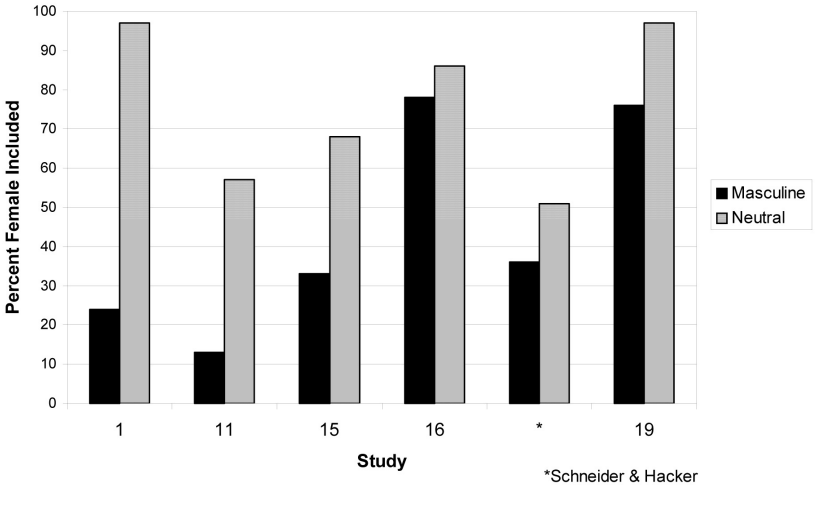


Figure 2. *Masculine/Neutral/Inclusive Comparison*

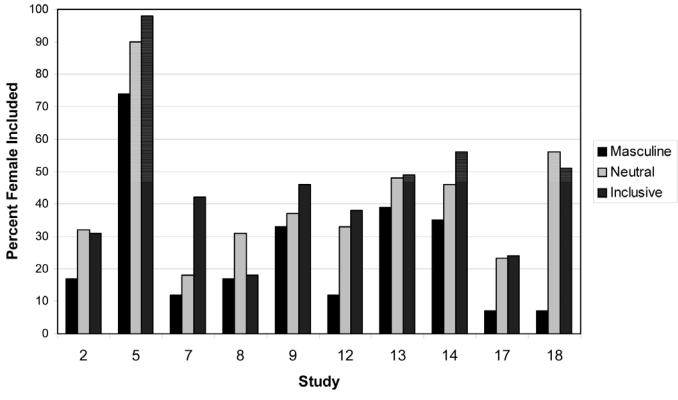
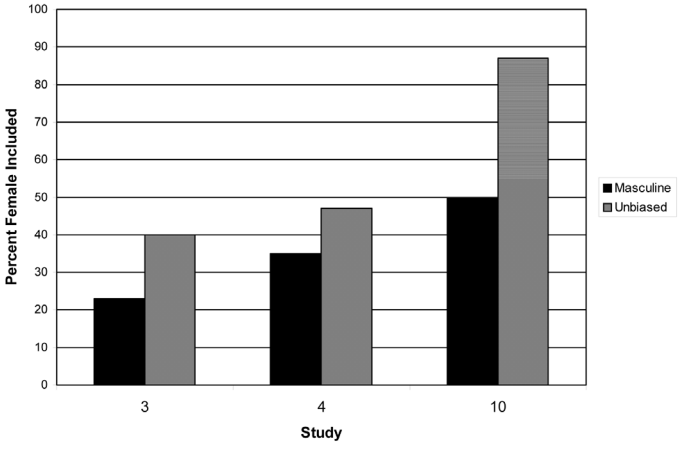


Figure 3. *Masculine vs. Unbiased Comparison*



The weighted *z* obtained for pronoun differences in inclusiveness, based on all 19 studies, was 9.10, $p < .00001$[10]. To avoid possible bias introduced by using more than one study by a single author, the weighted *z* was also calculated using only the first study listed by each author (eliminating 4, 6, 8, 11, 13); the weighted *z* for these 14 studies is 8.35, $p < .00001$. The conclusion from the aggregated data is that the masculine form is not a true generic, that is, it does not successfully reference both female and male, but overwhelmingly invokes a male-oriented response. Additional questions about the understanding of the masculine and other generics are addressed in the following sections.

[10] The probabilities found in the various tests in this article (sections 3 and 4) are unusually small; if the actual probabilities from statistical tables were given, they would begin with five to nine zeros. This is true in part because of large sample sizes.

4. Related questions of inclusiveness

4.1. Participant sex difference in inclusiveness

In most of the studies cited above, a difference by participant sex is reported. In all these studies finding significant difference, females were found to interpret generics more generically, that is, to give more female-inclusive responses, than did males, on at least one, and usually all, measures (although generally not giving more female- than male-inclusive responses). These results were also examined using the weighted Stouffer combined test.

The resultant z value, using all 19 studies, is 4.92, p <.00001. Again, the meta-analytic test was repeated with only one study from each author, with a resultant z of 4.33, p <º.00001. The sex difference finding is thus found to be highly robust.

Nilsen (1977) offers an explanation for the difference based on boys' retaining the childhood tendency to use the pronoun of their own sex when referring to the subject of a generic sex-neutral sentence. Then as boys get older, they continue to use pronouns as they did when young, but as girls get older, their speech is supposed to become "an inverse picture of their naturally developed childhood speech" (Nilsen, 1977: 178). In other words, boys and men may not utilize or understand the generic meaning of the MG because they have never had to learn to; girls and women, on the other hand, have had to learn to use the MG, and to believe in its inclusiveness, if they are to think of themselves and their sex as existing. Silveira (1980) has a discussion of these issues and the "people = self bias", extending it to adults as well.

Some obvious reasons for the sex difference in masculine generic use and comprehension which do not depend on this developmental pattern are same-sex identification and/or self-interest on the part of males (since use of the MG contributes, however unconsciously, to the perpetuation of societal male dominance).

4.2. Inclusiveness of sex-neutral vs. sex-inclusive forms

Some of the studies cited above have used both the I and the N forms of unbiased pronoun; does either one turn out to be better at eliciting inclusive responses? While Figure 3 suggests that I forms are superior in fostering inclusive thinking, overall the findings were mixed, with some studies showing a significant difference, others showing no significant difference, and most either not reporting the two forms separately, or not reporting any comparison tests. Because of insufficient information, the differences could not be subjected to meta-analysis. Further study, with more explicit tests and reporting, may show whether or under what circumstances either form may be superior.

4.3. People = male

The pervasiveness of masculine-biased response under any conditions was noted earlier. Silveira (1980) has named this general tendency to think of people as male the "people = male bias", a subcategory of the more general "generic = specific" bias found in studies of categorization and typicality (Rosch & Mervis, 1975). In this case, Silveira noted, when *man* was used, "woman" was not a typical member of the category. A people = male bias is not only operative, but is strengthened when the MG is used.

Many other studies support the prevalence of the people = male bias, in their findings of predominant male responses regardless of generic stimulus condition. For example, Martyna (1978) found that people completed N sentence fragments with *he* 65% of the time, and generated masculine names for U generic sentences 62-67% of the time. Wise & Rafferty (1982) found that when given N stimuli, their participants reported thinking of males over four times as often as of females (adult and child data summed). Schaeffer (1982) reported that regardless of whether the stimulus had M or N wording, boys were thought of approximately five times more often than girls. Gelb (1989) asked children to tell stories about a baby or animal in three sex-indeterminate pictures they viewed; he found that the children attributed maleness to each of the pictures, with 88.1% of the stories overall being about males.

Hamilton (1991) tested the people = male bias explicitly and in both directions. In Study 1, she had college participants complete sentences that required referring to either a male or female character, and participants had the option of choosing a neutral referent (e.g., *person* or *individual*) or a gender-specific referent (e.g., *man* or *woman*). She found that five times as many participants used a neutral referent for the male character, and used a gender specific referent for the female character, than the reverse. Hamilton's Study 2 reinterpreted the data of Hamilton & Henley (1982). Here she found that, when given U pronouns, male respondents reported significantly more male than female images; female respondents reported slightly more male than female images, but that difference was not significant. In Study 3, she showed that college participants were more likely to imagine that a "typical person" was male than female. These findings support not only the hypothesis that a person is more likely believed to be male than female, but also the notion that a male is more likely than a female to be considered a "person".

4.4. Age differences and understanding of the inclusiveness rule

Some studies have used adults, some children; are adults less prone to the MG error than children? The various studies summarized in this article suggest that in practice neither children nor adults apply the MG rule well in comprehension. To look more closely at the age-related difference in generic comprehension, we ran Stouffer's weighted combined test separately (there is no direct test of age difference

in any studies) on studies of (a) elementary and junior high participants (1, 5, 6, 8, 15, 18); and (b) high school, college, and adult participants (2, 3, 4, 9, 10, 11, 12, 13, 14, 16, 17, 19)[11]. The z value obtained with the younger participants was 6.50, p <.00001; that obtained with the older participants was 6.03, p <.00001. Thus the effect of the MG to influence masculine interpretation is strongly active in both younger and older participants[12].

But even if they do not apply the MG rule well, we might expect that when asked about it directly, adults and many middle- and older-aged children could explain it. Hyde (1984), in her two experiments, asked a series of questions to ascertain her participants' explicit knowledge of the grammatical rule for the MG, and found that 28% of first graders, 15-32% of third graders, 34-42% of fifth graders, and 90% of college students understood that *he* can mean both males and females.

Sniezek & Jazwinski (1986) had similar findings with adults. After completion of their experiments, these researchers queried college students on their understanding of the MG; they report that in one study, seven out of 56 (12.5%) participants were excluded because they did not understand it, and in another, data of nine participants out of 85 (10.6%) were excluded for the same reason. These figures are roughly in keeping with Hyde's finding of 90% understanding among college students.

Gastil (1990), however, did not find such high comprehension. He reports the results of an informal questionnaire and interview given to 44 of his 93 college student participants, assessing their comprehension of the inclusiveness rule (Footnote 2). From these responses, one third clearly understood it, a quarter thought it was outdated and/or dumb (presumably these participants also understood it), a quarter were not sure they understood it at all, and the rest (about 17%) did not provide clear answers. These figures suggest a comprehending percentage in adults of somewhat less than the 90% Hyde found, perhaps as low as Gastil's 58%[13].

4.5. Overview of other questions of inclusiveness

Summarizing the findings on these other questions of inclusiveness, then:

a) Females have quite consistently been found to respond to the MG more inclusively than have males. Possible reasons for the sex difference may be a

---

[11] Hyde's (1984) Experiment 1 could not be used because older and younger participants' data were not reported separately.

[12] The sex difference probability might have been smaller if six of the 19 experiments did not have missing probabilities, whose values were set at .5.

[13] It is both frustrating and rather interesting that none of these articles analyze figures for understanding of the inclusiveness rule by participant sex. On the basis of existing findings, we would predict that men's explicit understanding of the MG, like their implicit understanding, would be less inclusive of females than womens'.

continuation of childhood language patterns, same-sex identification, and self-interest among males, who have less reason than females to develop a generic comprehension, within the larger context of male dominance.

b) Sex-inclusive forms like *girls and boys*, *he and she* may produce more female-including responses than sex-neutral forms like *people* and *they*, though the evidence is mixed. However, those sex-inclusive forms are the forms many criticize as "awkward", "ungainly", and "cumbersome" (see Henley, 1987).

c) Part of the predominantly male-centered response to the MG is due to a general tendency to think of people as male to begin with; however, MGs increase that tendency.

d) The poor inclusiveness of responses to the MG may be a result of relatively poor understanding of the rule mandating the masculine for the generic, even in older children and, to some extent, in adults.

5. Conclusion

There seems to be no need for more research on whether masculine forms can serve as true generics in English. The existing research is unequivocal on this question: they cannot. Both sex-inclusive and sex-neutral forms aid people's generic comprehension more than MG. Several other questions have been clarified by systematic examination of the literature as well: females and males are both vulnerable to the influence of the MG, but males to a far greater extent than females; adults are highly vulnerable to it as well as are children; though adults' explicit knowledge of the generic rule understandably surpasses children's, it is surprising how many at the college level are unable to explain it satisfactorily.

Another major conclusion to draw from this literature is that MG comprehension in English is quite complex, affected by an underlying tendency to think of typical humans as males (and vice versa), and by a general tendency in language processing to match the (implied) gender of nouns and pronouns, among other factors. No doubt such factors are even more complex in languages with more pervasive gender structure.

These factors may pale, however, beside the global context of male hegemony which nourishes and in turn is supported by the prescription of masculine generics. MacKay (1980) has pointed out that the MG bears a close resemblance to highly effective propaganda techniques, with the advantage of frequency, covertness, early age of acquisition, association with high-prestige sources, and indirectness. Being required to speak of people as males comes to seem natural, immutable and preferable; and, being covert, is not easily perceived or challenged.

Is this topic still important enough to pursue? Is it trivial? Despite the persistence of the MG in some writing, does the younger generation really use it?

Some research has found the persistence of the MG among young people, who might be expected to be open to nonsexist language. Meyers (1990) found that slightly more college students used the MG in written proposals than used singular *they*. Although Noll (1999) found, in analyzing samples of editorial pages and college students' speech and written essays, that the students used singular "they" frequently, she also found editorial writers (influential language users), were still more likely to use the MG than singular "they".

As for the supposed triviality of the MG: its persistence over centuries, the vehemence with which it is defended (Martyna, 1980a), and its demonstrated links to person perception and subjective definitions of who is a person suggest otherwise. A testimony to the continuing importance accorded the study of generic forms in the field of language and gender is the recent appearance of books such as the projected three-volume anthology *Gender across languages*, which examines the manifestation of gender in a broad variety of world languages (Hellinger & Bussmann, 2001), and *Pronoun envy* which examines literary texts "to see how far the gender system can be pushed in English and French" (Livia, 2001: 10).

Despite all the research carried out on the MG, more remains to be done. We have already noted the incomplete state of knowledge regarding the comparative ability of different unbiased forms to assist generic interpretation. There are numerous studies showing various detrimental effects of MG use, especially on judgments of suitability for employment, but research showing how to overcome this problem would be welcomed. There have been some studies of change in usage, mostly observational; more experimental studies on the efficacy of different methods of changing MG use could be very helpful. We need research on the neglected area of intra- and interpersonal effects, such as on the emotional responses of girls (and boys) learning the MG rule. Cognitive scientists have realized that the study of MG processing has interesting implications for certain language processing issues, and systematic research in this area may be expected to expand.

Different and additional issues arise in languages other than English, such as what changes are appropriate given the structure of gender in the language; for example, the formation of agentives/occupational terms in several languages with grammatical gender has occupied much scholarly attention. No doubt new questions will be raised as research progresses; it may be expected to progress until females can find themselves as linguistically central as males are.

Bibliographical references

Banaji, M.R. & C.D. Hardin (1996). "Automatic stereotyping". *Psychological Science* 7, 136-41.

Baron, D.E. (1981). "The epicene pronoun: The word that failed". *American Speech* 56, 83-97.

Baron, D.E. (1986). *Grammar and gender*. New Haven, CT: Yale University Press.

Bodine, A. (1975). "Androcentrism in prescriptive grammar". *Language in Society* 4, 129-46.

Bucholtz, M., A.C. Liang & L.A. Sutton (eds.) (1999). *Reinventing identities: The gendered self in discourse*. New York: Oxford University Press.

Cameron, D. (1998). "Gender, language, and discourse: A review essay". *Signs* 23, 945-73.

Carreiras, M., A. Garnham, J. Oakhill & K. Cain (1996). "The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish". *Quarterly Journal of Experimental Psychology: Human Experimental Psychology* 49A, 639-63.

Cole, C.M., F.A. Hill & L.J. Dayley (1983). "Do masculine pronouns used generically lead to thoughts of men?". *Sex Roles* 9, 737-49.

Cooper, R.L. (1984). "The avoidance of androcentric generics". *International Journal of the Sociology of Language* 50, 5-20.

Crawford, M. & L. English (1984). "Generic versus specific inclusion of women in language: Effects on recall". *Journal of Psycholinguistic Research* 13, 373-81.

DeStefano, J.S. (1978/79). "Sex differences in language: A cross-national perspective with emphasis on English". *Language Sciences* 1, 316-24.

DeStefano, J.S., M.W. Kuhner & H.B. Pepinsky (1978). "An investigation of referents of selected sex-indefinite terms in English". Paper presented at the *Ninth World Congress of Sociology*, Uppsala.

Doleschal, U. & S. Schmid (2001). "Doing gender in Russian: Structure and perspective". In M. Hellinger & H. Bussmann (eds.), *Gender across languages: The linguistic representation of women and men* (Vol. 1). Amsterdam & Philadelphia: John Benjamins, 253-82.

Eberhart, O.M.Y. (1977). "Elementary students' understanding of certain masculine and neutral generic nouns" (Doctoral dissertation, Kansas State University, 1976). *Dissertation Abstracts International* 37, 4113A-14A.

Ehrlich, S. & R. King (1992). "Gender-based language reform and the social construction of meaning". *Discourse and Society* 3, 151-66.

Foertsch, J. & M.A. Gernsbacher (1997). "In search of gender neutrality: Is singular *they* a cognitively efficient substitute for generic *he*?". *Psychological Science* 8, 106-11.

Friederici, A.D. & T. Jacobsen (1999). "Processing grammatical gender during language comprehension". *Journal of Psycholinguistic Research* 28, 467-84.

Gastil, J. (1990). "Generic pronouns and sexist language: The oxymoronic character of masculine generics". *Sex Roles* 23, 629-43.

Gathercole, V.C. (1989). "The acquisition of sex-neutral uses of masculine forms in English and Spanish". *Applied Psycholinguistics* 10, 401-27.

Gelb, S. (1989). "Language and the problem of male salience in early childhood classroom environments". *Early Childhood Research Quarterly* 4, 205-15.

Hamilton, M.C. (1988). "Using masculine generics: Does generic he increase male bias in user's imagery?". *Sex Roles* 19*,* 785-99.

Hamilton, M.C. (1991). "Masculine bias in the attribution of personhood: People = male, male = people". *Psychology of Women Quarterly* 15*,* 393-402.

Hamilton, M.C. (1997). "The huwom race: Sexist language as a tool of dominance". In F. Braun & U. Pasero (eds.), *Kommunikation von geschlecht* [*Communication of Gender*]. Kiel: Zentrum für interdisziplinäre Fravenforschung der Christian-Albrechts, Universität zu Kiel, 147-63.

Hamilton, M.C. & N.M. Henley (1982). "Detrimental consequences of generic masculine usage". Paper presented at the meeting of the *Western Psychological Association*, Sacramento, CA.

Harrison, L. (1975). "Cro-Magnon woman –In eclipse". *Science Teacher* 42(4), 9-11.

Harrison, L. & R.N. Passero (1975). "Sexism in the language of elementary school textbooks". *Science and Children* 12(4), 22-25.

Heise, E. (2000). "Sind Frauen mitgemeint? Eine empirische Untersuchung zum Verstaendnis des generischen Maskulinums und seiner Alternativen" ["Are women included? An empirical study of the generic masculine and its alternatives"]. *Sprache & Kognition* 19, 3-13.

Hellinger, M. & H. Bussmann (2001). *Gender across languages: The linguistic representation of women and men* (Vol. 1). Amsterdam & Philadelphia: John Benjamins.

Henley, N.M. (1987). "This new species that speaks a new language: On sexism in language and language change". In J. Penfield (ed.), *Women and language in transition*. Albany, NY: State University of New York Press, 3-27.

Henley, N.M. (1989). "Molehill or mountain? What we know and don't know about the masculine as generic in English". In M. Crawford & M. Gentry (eds.), *Gender and thought: Psychological perspectives*. New York: Springer-Verlag, 59-78.

Henley, N.M., M.C. Hamilton & B. Thorne (1984). "Womanspeak and manspeak: Sex differences and sexism in communication, verbal and non-verbal". In A. Sergent (ed.), *Beyond sex roles*. Palo Alto, CA: West Publishing, 201-18.

Hyde, J.S. (1984). "Children's understanding of sexist language". *Developmental Psychology* 20, 697-706.

Jaworski, A. (1989). "On gender and sex in Polish". *International Journal of the Sociology of Language* 78, 83-92.

Khosroshahi, F. (1989). "Penguins don't care, but women do: A social identity analysis of a Whorfian problem". *Language in Society* 18, 505-25.

Kidd, V. (1971). "A study of images produced through the use of the male pronoun as the generic". *Moments in Contemporary Rhetoric and Communication*

1(2), 25-30.
Lakoff, R. (1975). *Language and woman's place*. New York: Harper & Row.
Livia, A. (2001). *Pronoun envy: Literary uses of linguistic gender*. New York: Oxford University Press.
Lozano, I. (1995). *Lenguaje femenino, lenguaje masculino: ¿Condiciona nuestro sexo la forma de hablar?*. Madrid : Minerva.
MacKay, D.G. (1980). "Psychology, prescriptive grammar and the pronoun problem". *American Psychologist* 35*,* 444-49.
MacKay, D.G. & D.C. Fulkerson (1979). "On the comprehension and production of pronouns". *Journal of Verbal Learning and Verbal Behavior* 18*,* 661-73.
Martyna, W. (1978). "Using and understanding the generic masculine: A social-psychological approach to language and the sexes" (Doctoral dissertation, Stanford University, 1978). *Dissertation Abstracts International* 39*,* 3050B.
Martyna, W. (1980a). "Beyond the he/man approach: The case for nonsexist language". In B. Thorne, C. Kramarae & N.M. Henley (eds.), *Language, gender and society*. Rowley, MA: Newbury House Publishers, 25-37.
Martyna, W. (1980b). "The psychology of the generic masculine". In S. McConnell-Ginet, R. Borker & N. Furman (eds.), *Women and language in literature and society*. New York: Praeger, 69-78.
Meyers, M.W. (1990). "Current generic pronoun usage: An empirical study". *American Speech* 65, 228-37.
Mosteller, F.M. & R.R. Bush (1954). "Selected quantitative techniques". In G. Lindzey (ed.), *Handbook of social psychology* (Vol. 1). Cambridge, MA: Addison-Wesley, 289-334.
Moulton, J., G.M. Robinson & C. Elias (1978). "Sex bias in language use: "Neutral" pronouns that aren't". *American Psychologist* 33*,* 1032-36.
Ng, S.H. (1988). "Masculine generics and sexism in language". Final report to the *Social Sciences Research Fund Committee*, Wellington.
Ng, S.H. (1990). "Androcentric coding of *man* and *his* in memory by language users". *Journal of Experimental Social Psychology* 26, 455-64.
Nilsen, A.P. (1977). "Sexism in children's books and elementary teaching materials". In A.P. Nilsen, H. Bosmajian, H.L. Gershuny & J.P. Stanley (eds.), *Sexism and language*. Urbana, IL: National Council of Teachers of English, 161-79.
Noll, J.A. (1999). "Studies of epicene reference" (Doctoral dissertation, University of South Florida, 1999). *Dissertation Abstracts International* 60(3), 1324B.
Pauwels, A. (1998). *Women changing language*. London: Longman.
Romaine, S. (1999). *Communicating gender*. Mahwah, N.J.: Lawrence Erlbaum Associates.
Rosch, E. & C.B. Mervis (1975). "Family resemblances: Studies in the internal structure of categories". *Cognitive Psychology* 7*,* 573-605.

Rosenthal, R. (1991). *Meta-analytic procedures for social research*. (Rev. ed.), Thousand Oaks, CA: Sage Publications.

Rothermund, K. (1998). "Automatische geschlechtsspezifische Assoziationen beim Lesen von Texten mit geschlechtseindeutigen und generisch maskulinen Textsubjekten" ["Automatic gender-specific associations during the reading of texts containing gender-specific and masculine generic text-subjects"]. *Sprache & Kognition* 17, 183-98.

Schaeffer, N.C. (1982). "A General Social Survey experiment in generic words". *Public Opinion Quarterly* 46, 572-81.

Schneider, M.J. & S.L. Hacker (1973). "Sex role imagery and the use of the generic "man" in introductory texts". *American Sociologist* 8(8), 12-18.

Schriefers, H. & J. Jescheniak (1999). "Representation and processing of grammatical gender in language production: A review". *Journal of Psycholinguistic Research* 28, 575-600.

Shimanoff, S.B. (1977). "Man = human: Empirical support for the Whorfian hypothesis". *Bulletin: Women's Studies in Communication* 1(2), 21-25.

Silveira, J. (1978). "Women on the fringes: Generic masculine words and their relation to thinking". Unpublished manuscript.

Silveira, J. (1980). "Generic masculine words and thinking". *Women's Studies International Quarterly* 3, 165-78.

Sniezek, J.A. & C.H. Jazwinski (1986). "Gender bias in English: In search of fair language". *Journal of Applied Social Psychology* 16*,* 642-62.

Stouffer, S.A., E.A. Suchman, L.C. De Vinney, S.A. Star & R.M. Williams (1949). *The American soldier: Adjustment during life* (Vol. 1). Princeton, NJ: Princeton University Press.

Switzer, J.Y. (1990). "The impact of generic word choices: An empirical investigation of age- and sex-related differences". *Sex Roles* 22, 69-82.

Thorne, B. & N.M. Henley (eds.) (1975). *Language and sex: Difference and dominance*. Rowley, MA: Newbury House Publishers.

Wilson, E. & S.H. Ng, (1988). "Sex bias in visual images evoked by generics: A New Zealand study". *Sex Roles* 18(3/4), 159-68.

Wilson, L.C. (1978). "Teachers' inclusions of males and females in generic nouns". *Research in the Teaching of English* 12*,* 155-61.

Wise, E. & J. Rafferty (1982). "Sex bias and language". *Sex Roles* 8, 1189-96.

Wolf, F.M. (1986). *Meta-analysis: Quantitative methods for research synthesis.* Beverly Hills, CA: Sage Publications.