GRAMMATICAL GENDER
AND GENDER IDENTITY DEVELOPMENT:
Cross Cultural and Cross Lingual Implications

Benjamin Beit-Hallahmi, J. C. Catford, Ralph E. Cooley, Cecelia Yoder Dull, Alexander Z. Guiora, Maria Raluszny

The University of Michigan Personality and Language Behavior Research Project, Ann Arbor, Mich.

Conceptual problems confronting research relating language and cognition are examined, focusing on the relationship of linguistic gender to gender identity as a personality construct. A research strategy, combining psycholinguistic considerations and developmental concepts is proposed, to provide an empirical test of the influence of native language on individual maturation.

The basic question of the relationship between language and cognition has remained a difficult but intriguing problem, and efforts to demonstrate a relationship between the two have led to some controversy and confusion. That there is some relation between language and cognitive processes seems obvious, but an examination of the literature suggests caution to those who undertake research in this area. Our concern in this article is the general relationship between linguistic and non-linguistic behavior and the specific relationship of linguistic gender to gender identity as a personality construct. Social-psychological, psycholinguistic, and developmental perspectives are combined in discussing research implications in this area.

Two conceptual problems confront research relating language and cognition. First, evidence of linguistic differences alone is not sufficient for making inferences about cognitive differences; second, care must be taken in specifying the level of the type of cognitive activity being evaluated (*e.g.,* basic intellectual processes, perception, etc.) before offering conclusions of any kind. One possible starting point for this discussion can be found in the linguistic relativity hypothesis, most commonly associated with the names of Sapir and Whorf. The idea

of linguistic relativity [4] certainly goes back further than Whorf: it is, for instance, a natural consequence of the widely accepted view, propounded by de Saussure,[5] that any item in a linguistic system acquires a "value" from its opposition to other items in the system: it was, moreover, expressed before Whorf in quite Whorf-like terms in Lee.[13] Nevertheless, the hypothesis of linguistic relativity was, perhaps, most explicitly stated by Whorf, whose view is summed up by Carroll,[3] in these words:

The structure of a human being's language influences the manner in which he understands reality and behaves with respect to it.

In other words, it is predicted that the particular language that an individual uses affects the way he understands his environment. Different languages present different pictures of the words; a language may be said to analyze reality, and the relativity of this analysis is demonstrated by the different ways in which different languages reflect or codify identical aspects of reality, such as colors or sexes. The Sapir-Whorf hypothesis is introduced here since it has often been used to conceptualize the relationship between language and cognition and the impact of linguistic relativity on the latter. Brown and Lenneberg [2] spell out two propositions involved in the Sapir-Whorf hypothesis. The first suggests that different linguistic communities perceive reality in different ways, and the second suggests that the language spoken in a community takes part in shaping the cognitive structure of individuals in that community. Whorf [18] supplied numerous examples of language relativity, most of which are related to environmental and social differences. One may, however, go one step further and distinguish two kinds of linguistic relativity, lexical and grammatical. Lexical relativity (which in most of Whorf's examples means differences in distinctions among aspects of reality) may be related to environmental or social determinants.

Fishman [7] suggested that the Sapir-Whorf hypothesis could be tested at four separate levels of complexity and generality. He refers to the analysis done in terms of lexical specificity as an analysis at the linguistic codifiability level. This kind of analysis uses only linguistic data, and according to Fishman is the lowest level at which the hypothesis can be tested. Analysis at this purely linguistic level can provide us with interesting illustrations, but adds only little to our understanding of higher levels, because linguistic differences are not related, at this level of analysis, to non-linguistic behaviors. The level of analysis becomes more complex and crucial to the testing of the hypothesis, when linguistic and non-linguistic behaviors are correlated. As examples of experimental studies testing the Sapir-Whorf hypothesis at the level of relating linguistic and non-linguistic behavior, we can mention the work of Brown and Lenneberg [2] and Niekawa-Howard.[15, 16] These studies, relating language characteristics to perception, provide positive indications of the validity of the hypothesis.

Among the aspects of language that give rise to linguistic relativity, Brown [1] distinguishes between lexical features, such as the specific terms for three kinds of snow in the Eskimo language, and grammatical features, such as gender, number, tense, etc. Some grammatical distinctions, such as gender-assignment, often appear to be arbitrary with respect

to reality: why sin is feminine in German but masculine in Russian cannot be discovered by any simple analysis of environmental or social determinants. Moreover, there are differences among languages in the extension of gender-marking. In many languages it applies only to nominal categories, such as pronouns and adjectives; in some, gender-marking also applies to verbs and may in certain cases be directly related to the sex of human participants in the speech situation: though biological differences of sex are universal, the extent to which such differences are obligatorily distinguished in grammar differs from language to language.

LINGUISTIC ANALYSIS AND PSYCHOLOGICAL IMPLICATIONS OF GENDER

Two linguistic categories relevant to this discussion are examined in the following analysis, namely person and gender. A universal characteristic of all occasions where languages are used is the necessary existence of two human participants in the language-event—a performer (the speaker or writer who generates the utterance) and an addressee (hearer or reader) and any utterance necessarily makes reference to either one of these human participants (performer or addressee), or else to some other participant in the event—human or nonhuman, present or absent. These three categories of participants are what traditional grammar calls persons: first person (performer), second person (addressee), and third person (other). All known languages oblige their speakers to distinguish these universal participant roles no matter in what form they are manifested in the situation on any particular occasion. In all languages there are grammatical rules calling for

the obligatory selection of particular forms—most commonly of pronouns and verbs—relating to at least the three persons defined above, and to some combinations of these.

Another very common, though not universal, characteristic of languages is the repartition of the nouns of a language into a number of different categories or classes. These classes are not always made evident by the forms of the nouns themselves but are marked by the phenomenon of "agreement" or "grammatical concord." This is the well-known grammatical category of gender. We are familiar with languages such as Latin, French, and German, where nouns fall into gender classes indicated by the forms of adjectives, pronouns, and articles with which they are syntactically related.

We refer to gender as a formal category of grammatical concord, and this is the way that linguists usually regard it. In the context of the present article, we must give attention to the semantic aspect of gender: to what extent do different gender classes correspond to non-linguistic categories—in particular to the biological category of sex? It is often supposed that the correlation with sex is quite close, hence the commonly used terms, masculine, feminine, and neuter. In some languages there is, indeed, such a relationship; but in no language is it very close. When we deal with languages having gender-systems consisting of a large number of classes (up to eight in certain languages of the Caucasus area, up to fifteen or so in Bantu languages), it is quite clear that we cannot be dealing only with sex. For example, the multigendered Caucasian languages all make a male versus female distinction with respect to humans: but the remain-

ing categories have nothing to do with sex and are hard to define. Even in more familiar languages, such as French and German, in which we have masculine and feminine and masculine, feminine, neuter, respectively, there is little correlation of gender with sex. There is some correlation, in the sense that most nouns referring to males are masculine and most nouns referring to females are feminine. But even this correlation is not absolute: as is well known, the French word for *sentinel* is feminine, and the German words for *woman* and *girl* are neuter.

In English, sex is of some importance only in the selection of the third person pronouns, he, she, and it; it figures nowhere else in the grammar. There are other languages in which sex plays absolutely no grammatical role at all: the Finno-Ugrian languages, of which Hungarian and Finnish are the best known examples, are a case in point. They contain, as do all languages, specific lexical items meaning *man, woman, boy, girl,* etc. in which the sex of the referent is part of its meaning: but sex plays no part at all in the obligatory selection of grammatical categories. On the other hand, there are some languages, such as the Semitic languages, in which sex plays a much greater role in grammar. In these languages, there is a two-term system of grammatical gender—all nouns are either masculine or feminine, with about the same small degree of correlation with sex as in French. Where the Semitic languages differ very significantly from the languages so far considered is that, in their pronoun systems, the sex-gender distinction is not confined to the third person. In both Arabic and Hebrew, pronouns of the second person (you) are differentiated by gender. This

is very important in the present context since, as we saw above, the second person refers to the *addressee*, who is one of the human participants in the speech-event. Consequently, in these languages, the speaker is obliged to respond differentially to the sex of a human being whenever he selects a second-person pronoun. In Arabic, this sex-determined gender-selection is confined to the singular; in Hebrew it extends also to the plural.

Among the languages we have mentioned, there are not only differences in pronoun systems but also differences with respect to verb-forms. In the Finno-Ugrian languages verb-forms mark person distinctions but have no trace of a distinction dependent on the sex of any participant. The English verb marks some person and number distinctions (am-are, is, was-were, go-goes) but has no distinction in verb-forms dependent on gender or sex. The Semitic languages once again stand out in this respect. In Arabic, gender (masculine vs. feminine) is marked in the second and third person singular of all tenses; in the third person the selection of masculine or feminine forms is of course determined by the gender-class of the subject, which may be a noun referring to anything at all, including any inanimate object, and hence is only very slightly correlated with sex. In the second person, however, the sex of the *human* addressee directly determines the form of the verb.

In Hebrew, the masculine versus feminine gender distinction is marked in the second and third person plural, as well as in the singular, of all finite verb-forms. In the present tense of verbs this distinction is extended to the first person as well. Thus, the Hebrew speaker, every time he uses a verb in

the present tense, is forced to be aware of his own sex, if he is referring to himself as performer (first person), and of the sex of the addressee, when he is using the second person. Consequently, in a Hebrew-speaking environment, a child must be constantly reminded of the distinction between the sexes in the way grownups of different sexes select gender forms in addressing the child and each other, and in the requirements of gender selection that they impose on him.

We see, then, that languages differ not only in the extent to which they employ grammatical gender—manifested in pronominal and sometimes verbal agreement in the third person—but also in the extent to which they oblige their speakers to take note of *sex*. In the Finno-Ugrian languages the sex of participants in the language event is never grammatically significant at all. In English it is only marginally so. In Semitic languages it becomes quite an important factor in determining selection of grammatical forms. We could sum this up by saying that the sex-determined grammatical "gender loading" of language varies from *zero* in languages like Finnish and Hungarian, through *very low* in English, to *very high* in Hebrew.

THE HYPOTHESIS

Some of the possible implications of gender for the perception of linguistic stimuli have been discussed by Ervin.[6] The question we are pursuing here is the relationship between gender structure (a linguistic construct) and individual gender identity (a psychological construct). If we go back to the Sapir-Whorf hypothesis, presented above, the most general question is: how will differences in gender loading of words affect the perceptions of reality in different cultures? In particular, what is the impact—if any—of varying degree of linguistic gender loading on the development of gender identity in the growing child? Empirically, the question is: will there be a correlation between the amount of linguistic emphasis on sex-determined gender and the average age of attaining gender identity in children, in a specific linguistic environment?

RESEARCH ON GENDER IDENTITY

The concept of gender identity, as used here, refers only to the individual's ability to categorize himself as a member of one sex and not the other. Attainment of gender identity is thus operationally defined as the demonstrated ability of the child to indicate his or her own sex. Furthermore, the emphasis here is on the early stages of gender identity formation, not its completion, nor the successful completion of "sexual identity" or "sex role." Nevertheless, the attainment of gender identity in our (limited) sense is a clear and distinct developmental milestone, which is perhaps the "essential foundation on which subsequent gender identity is built." [12]

Empirical data on the development of gender identity is based mainly on observational material. Reports in the psychoanalytic and developmental psychological literature [8, 9] identify the latter part of the third year of life as the period in which gender identity is attained. However, no systematic investigation of gender identity in children below 36 months of age, and no attempts to demonstrate the age range within which such gender identity may occur, have been reported in the literature. What has hampered systematic research on gender identity has been the lack of a valid and reliable instrument. The only

measure currently in wide use is Ge-sell's [9] procedure, which consists of ask-ing the direct question: "Are you a little boy or a little girl?" Despite its popu-larity and its wide use by pediatricians and other clinicians, the method is open to criticism. The child has a chance of being correct half of the time, while his failure to respond correctly may be caused by a variety of uncontrolled variables.

In response to the need for a suitable measure, the Michigan Gender Identity Test (MIGIT) has been developed.[17] The aim in developing this instrument was to provide a valid, reliable measure that could be used in a variety of experi-mental studies. In designing an instru-ment for measuring gender identity, two somewhat separate areas of concern served as major guidelines: the very young age of our test population and the cross-cultural nature of the study in which the test would eventually be utilized. In addition it was felt that such a test would be of optimal usefulness if it could show not only the presence or absence of gender identification but also demonstrate that gender identity, as a developmental process is related to the basic cognitive functions of classi-fication and concept formation.

Since the age range of the subjects we proposed to test was to be 16 months–48 months, it was essential in construct-ing the test to consider the capacities of the youngest subjects. What was re-quired was a brief, easily administered, gamelike procedure that would place minimal demands on the attention span, activity level, and social maturity of our young subjects. In addition, since the proposed age range would also involve a wide variation in level of language development, it was necessary to set up the test so that the responses could be nonverbal or verbal. Thus the less lin-guistically mature or perhaps simply bashful or reluctant child would be more accessible for fair assessment. Devising a "culture-free" test also meant choosing stimuli that would be familiar to chil-dren in the chosen cultures. The test we have developed involves assessing the child's performance on a series of pic-ture sorting tasks. The pictures include familiar non-human objects (dogs and balls) and young human beings (boys and girls). In a playlike fashion, the child is asked to sort the pictures into groups. In the case of the boys and girls, the sorting is of course according to sex. At the end, the child is asked to place his own picture with the appropriate sex-group and to respond to the Gesell question as described above. Successful completion of the MIGIT involves sort-ing all pictures correctly and then re-sponding to one's own picture in terms of gender. The results of research em-ploying the MIGIT so far show that the test is usable and produces meaningful results and also that the attainment of gender identity is a developmental pro-cess. The MIGIT has been successfully used with a group of 105 American chil-dren, aged between 16 and 51 months. In this study the developmental curve of the attainment of gender identity was found to rise sharply in the third year of life. Age-group analysis showed that, in this sample, 27% of those aged 28 months, 50% of those aged 30 months, and 100% of those aged 37 months completed the MIGIT successfully.

In comparisons between boys and girls by age group, no significant differ-ences were found. An analysis of error patterns on the MIGIT gave support to the notion that gender identity is a de-

velopmental process, related to general cognitive growth, and that test performance reflects this in an orderly and progressive manner. Since there are indications that the attainment of gender identity is indeed a developmental process, it may be of interest to look at the possible antecedents and correlates, and to determine what are the environmental and individual variables that may be related to variations in this process. Two possible variables of interest are *general level of maturation* and *general level of language development*. Recent contributions to the literature on gender identity development have supported recognition of the role of both cognitive abilities [11] and differential learning experiences [14] in this process, thus modifying the traditional viewpoint, which emphasized biological and maturational factors. We may hypothesize that family constellation, socioeconomic level, and cultural factors are all likely to be related to the attainment of gender identity in the individual. Since we assume a psycholinguistic orientation, we here pose the specific question: do children who grow up in a language culture that knows no gender distinction lag, in the attainment of gender identity, behind children who grow up in a language culture that has extensive sex-determined gender distinctions marked in verbs, pronouns, and adjectives?

The way in which we propose to answer this question is by administering the MIGIT to groups of children whose native languages differ in terms of gender loading, and who are otherwise (social class, sibling position) as similar as possible. If, for example, we administer the MIGIT to a sample of Israeli children, whose native language is Hebrew, and then to a sample of Finnish children, whose native language is Finnish, with the two samples closely matched in respect to social class and sibling position, the only major variable differentiating the two samples will be language. What we want to determine is whether the high gender loading in Hebrew, discussed above, and the absence of any gender loading in Finnish, will affect the timing of individual gender self-identification in these two populations. Our results should be of some significance regardless of the specific answer to this question. If comparative data show no difference between language groups differing in gender loading, one might conclude that the development of gender identity is largely uninfluenced by the powerful cultural factor of language. On the other hand, if an earlier achievement of gender identity co-exists with a language high in gender-loading, one might conclude that language as an environmental factor significantly influences the timing of what is, in part, a maturational process.

SUMMARY

The highest level of applying and testing the Sapir-Whorf hypothesis, according to Fishman,[7] is that of combining grammatical characteristics and nonlinguistic data. This is the level at which we are suggesting some interrelationships of language and individual personality development.

The rationale proposed here aims at a crucial test, which combines psycholinguistic notions with a developmental conception of gender identity attainment. An empirical test of these ideas will have an important bearing on both the notion of gender identity attainment as a developmental process, and on the psycholinguistic influence hypothesis.

The answer to the "psycholinguistic" question (*i.e.,* Does gender loading in one's native language have an influence on the timing of the attainment of gender identification?) is significant in understanding the general relationship between social determinants and individual maturation.

REFERENCES

1. BROWN, R. 1957. Linguistic determinism and the part of speech. J. Abnorm. Soc. Psychol. 55:1–5.
2. BROWN, R. AND LENNEBERG, E. 1954. A study in language and cognition. J. Abnorm. Soc. Psychol. 49:454–462.
3. CARROLL, J. 1956. Introduction. *In* Language, Thought, and Reality. Holt, New York.
4. CATFORD, J. 1969. Learning language in the field: problems of linguistic relativity. Mod. Lang. J. 53:310–317.
5. DE SAUSSURE, F. 1916. Cours de Linguistique Generale. Paris.
6. ERVIN, S. 1962. The connotations of gender. Word 18:249–261.
7. FISHMAN, J. 1960. A systematization of the Whorfian hypothesis. Behav. Sci. 5:323–339.
8. FRAIBERG, S. 1959. The Magic Years. Charles Scribner's, New York.
9. GESELL, A. 1940. The First Five Years of Life. Harper and Brothers, New York.
10. KAGAN, J. 1964. Acquisition and significance of sex-typing and sex-role identity. *In* Review of Child Development Research, M. Hoffman and L. Hoffman eds. Russell Sage Foundation, New York.
11. KAHLBERG, L. 1966. A cognitive-developmental analysis of children's sex role concepts and attitudes. *In* The Development of Sex Differences, E. Maccoby ed. Stanford University Press, Stanford, Calif.
12. KLEEMAN, J. 1971. The establishment of core gender identity in normal girls. I. (a) Introduction; (b) Development of ego and capacity to differentiate. Arch. Sex. Behav. 1:103–116.
13. LEE, D. 1938. Conceptual implications of an Indian language. Phil. Sci. 5:89–102.
14. LEVIN, S., BALISTRIERI, J., AND SCHUKIT, M. 1972. The development of sexual discrimination in children. J. Child Psychol. Psychiat. 13:47–53.
15. NIEKAWA-HOWARD, A. 1968. A psycholinguistic study of the Whorfian hypothesis based on the Japanese passive. Paper presented at Thirteenth Annual National Conference on Linguistics, New York.
16. NIEKAWA-HOWARD, A. 1968. A study of second language learning. Final Report, Project No. 3260, Contract No. OE-6-10-308, U.S. Department of Health, Education, and Welfare, Office of Education, Washington, D.C.
17. PALUSZNY, M., ET AL. 1973. Gender identity and its measurement in children. Comp. Psychiat. 14:3.
18. WHORF, B. 1956. Language, Thought and Reality. New York.