The Pragmatic Function of Modal Verbs in Scientific Papers

Atsuko K. Yamazaki
The Pragmatic Function of Modal Verbs in Scientific Papers

Atsuko K. Yamazaki*

Abstract

It has been pointed out that the precise use of a word can be critical in scientific papers. The epistemic use of modal verbs, such as must, may and might, can be important for scientists in this sense. The author has conducted a study to identify how much certainty scientists attribute to these modals used in scientific papers. Natural Scientists who spoke English as a native language were asked to indicate how much certainty they attributed to a concluding sentence containing modal and non-modal verbs in brief chemistry reports. Statistical analysis of the results showed that there was no significant difference between the certainties the scientists had attributed to might and may, although many English grammar books for non-native speakers mention that might is used to express less certainty compared to may. The results suggest that ESL/EFL teachers should be aware of such sociopragmatic functions of the modal verbs.

Key words: Modal verbs, Scientific writing, Pragmatic meaning, English Education

Introduction

English in technical writing is often pointed out as being different from “general” English in terms of the function of words (Applenquist 1981; Flick and Anderson 1980; Swales 1980). The language used in science is required to reflect the precise and objective nature of science, and scientists need to choose their words carefully. In particular, they need to pay attention to the function of words when they express their logic. When scientists use such modal verbs as must, may, could, might and should in research papers, the epistemic modality expressed with these auxiliary verbs can play an important role. When these modals are used to express certainty or possibility toward findings and hypotheses, their roles seem to be more significant, and the scientist, therefore, needs to carefully select an appropriate modal verb in order to convey his/her idea to the readers precisely.

English grammar books for scientists often advise that the use of such modal verbs as may, might and could should be avoided in scientific writing since they are considered to be the reflection of the scientist’s uncertainty about his/her findings or conclusions (Harada 1994; Yamamoto & Fukutake 1995). The imprecise use of a modal verb in a scientific paper could lead the reader towards making a wrong interpretation of the conclusion and could interfere with the purpose of scientific

* Part-time Instructor, Division of International and Interdisciplinary Studies, Tokyo University of Fisheries, Konan 4-chome, Minato-ku, Tokyo 108-8477, Japan.
papers, which is to state scientists' ideas and findings to their scientific community. Hyodo (1993) notes that the wrong use of modal verbs in scientific writing tends to make the credibility of argument unclear and may cause the reader to question the findings and conclusions discussed in the papers.

However, many scientists, especially those in the natural sciences, feel that it is necessary to use the modals to express inference in their papers, since nothing is 100% certain in science and inference is an important part of scientific research. By summarizing empirical studies on hedging in scientific discourse, Hyland (1994) points out the frequent occurrence of modal verbs in academic writing. Harada’s survey (1994) also shows that scientists often use modal verbs to express conjecture in their research papers. According to his survey on scientific writing, scientists who are native speakers of English use, on average, may and might more than seven times in ceramic science papers and more than three times in polymer chemistry papers. In her study of medical English abstracts, Salager-Meyer (1992) points out that modal verbs are frequently used in the recommendation, conclusion and data synthesis sections. In her result, the modals represent 45% of all the verb forms in the recommendation section and more than 20% in the conclusion section.

The use of modal verbs in “general” English has been studied by many linguists (Jespersen 1964; Palmer 1968; Quirk et al. 1985). Halliday (1985) gives a diagram to show probabilities expressed with propositions. The diagram rates must as carrying highest probability among modal verbs. In the use of modals to express certainty or probability, Celce-Murcia and Larsen-Freeman (1983) have established a hierarchy among modals. They rate could and might as having the lowest certainty, and may as having a higher degree of certainty than could and might, while they rank will as having the highest degree of certainty, and must as the second most certain, and then should as having less certainty than must. They also note that the degrees of probabilities expressed by these modals are not necessarily equi-distant: there is a smaller gap between may (also could) and might, but a bigger one exists between may and should in their probability scale when the modals are used affirmatively (Celce-Murcia and Larsen-Freeman 1983). Also, comparing may and might in terms of their functions, Quirk et al. (1985) note that might is used to express less certainty.

The use of these modals in scientific writing has been examined by Huddleston (1971). He concludes that may is often used to express uncertainty or possibility, might is an “unreal” counterpart to may regarding certainty/possibility, and must expresses something necessarily true (Huddleston, 1971). Although his analysis indicates the features of these modal verbs in scientific writing, it does not explain the semantic functions of these verbs in the discourse of scientific writing. The use of modal verbs as hedging devices in scientific writing has been discussed by many researchers (Adams-Smith 1984; Hyland 1994). Hyland examines the epistemic functions of modals, and claims that modals appear to be the typical devices to express hedging in scientific writing. He states that the modal auxiliaries are important means of allowing a scientist to adjust the degree of certainty about his/her claims and to build the writer-reader relationship that the writer wants to achieve.

In consequence, it is valuable to measure how the reader will interpret the degree of certainty/possibility attributed to each modal verb. Although the qualitative measurement of the epistemic
modality has been done for the modals in scientific writing (Salager-Meyer, 1992), how much certainty/possibility each modal verb carries has been rarely quantified according to their empirical uses in real scientific writing. In addition, quantifying the pragmatic meanings of modals with a scale of certainty would be readily accepted and understood by scientists, who are more familiar with quantitative representations than qualitative ones. An investigation to establish quantitative criteria for the epistemic uses of modal verbs in scientific writing seems important in this sense.

In this study, the author conducted a survey to examine the epistemic uses of the modal verbs such as must, may and might and their semantic functions in scientific writing. The study focused on the degree of certainty attributed to each modal verb by scientists. The statistically analyzed results of the survey will be presented, along with the interpretation of the results from pragmatic aspects of the modal verb use in the following sections. Based on the analysis, a hierarchical order of these modal verbs in terms of the degree of certainty regarding scientists' inferences will be suggested. Also, what the implications stemming from the results are for ESL/EFL teachers and scientists who are non-native speakers of English will be discussed.

Hypothesis

Three modal verbs, must, may and might were investigated in this study. In addition to these variables, non-modal verbs were included as a comparison. From Celce-Murcia and Larsen-Freeman’s (1983) hierarchy of certainties for these modals, and Huddleston’s (1971) analysis of the logical probability uses of must, may and might, the author has hypothesized that the degree of certainty which natural scientists who are native speakers of English attribute to the above modals in chemical research reports was as shown in Fig. 1.

Methods

i) Variables

The independent variable in this study was the use of modal verbs. This variable had four levels, the existence of must, may and might, and the nonexistence of modal verbs. In this study the control variable, a chemical research report, was defined as an abstract contained in Chemical Abstract, Vol. 92 (1980), which reports research results in chemistry. As another control variable in this study, a
“natural scientist who is a native speaker of English” was defined as a person who has been involved in chemical, physical or biological research for over three years, and whose first language is English. The dependent variable was the degree of certainty with which a natural scientist interprets modal and non-modal verbs in terms of percentages in chemical research reports.

ii) Subjects

The subjects were seventy one scientists (N = 71) who were native speakers of English involved in chemical, physical, biological or meteorological research at universities or companies in the states of California, Minnesota and New York. They were from twenty-seven to sixty-five years old, and had presented their own research papers at least once within the last three years. Fifty-nine of them (84.3%) were male. Forty-six of the subjects (65.7%) belonged to the American Chemical Society, twenty (28.6%) belonged to the American Association for Advancement of Science, eighteen (25.7%) were members of the America Association for Clinical Chemistry, and thirteen (18.9%) belonged to the American Meteorological Society. In addition, about 91% of the subjects gave a chemical journal and 40% gave a biological journal as one of three academic journals they read most often. About 19% of the subjects read academic journals in physics and 15.7% read meteorological journals. The academic journal read most often by the subjects (74.3%) was the Journal of American Chemical Society.

iii) Procedures

All the subjects were given a questionnaire which contained six different brief passages of chemical research reports. They were asked to show how much certainty they thought the author of each paragraph had toward his/her findings stated in the last sentence of the paragraph. The certainty was indicated by placing an “X” on a percentage scale ranging from 0 to 100%.

The questionnaires were presented in four formats (Format I through IV) of six paragraphs each. The first and the sixth paragraphs were exactly the same in all four formats. They were added to the questionnaires as “distractor” paragraphs. The second, third, fourth and fifth paragraphs were the same in all formats, except for the last sentence. The last sentence of each paragraph had a different modal or non-modal verb in each format. From the second to the fifth paragraph, the sentences did not contain any words which indicated certainty except for the modal verbs. Eighteen subjects (25.4%) responded to Format I, and fifteen subjects (21.1%) filled out a questionnaire in Format II. Eighteen subjects (25.4%) answered Format III and twenty subjects (28.2%) did Format IV. The combinations of the paragraphs and the verbs in each format is shown in Table 1, and an example questionnaire (Format I) is shown in Appendix I.
The Pragmatic Function of Modal Verbs in Scientific Papers

The six paragraphs were taken from the concluding parts of abstracts in *Chemical Abstracts*, Vol. 92 (1980) and modified to consist of three to five sentences for this study. These abstracts summarize research papers published in 1980 and cover the fields of biochemistry, organic chemistry, kinetic chemistry, analytic chemistry and electrochemistry. The paragraphs were the revisions of the ones tested in a pilot study conducted in 1983 (Yamazaki 1997) using the same procedure. The numbers and pages of the original abstracts in *Chemical Abstracts* are listed in Appendix II.

It should be noted that the population for this study was narrowly restricted to the population of researchers engaged in chemical research, in order to increase the validity of the study and the control confounding selection of variables. By the same token, the abstracts from chemistry papers were used as scientific reports.

**Data Analysis**

The Number Cruncher Statistical System V. 5.6 (Hintze 1990) was employed for data analysis. In order to examine whether the degree of certainty given to each modal verb or non-modal verbs differed according to the contents of the four paragraphs (paragraphs 2 through 5), a one-way analysis of variance (ANOVA) was computed among the paragraphs for each modal verb and the non-modal verbs. Then, the differences in the degree of certainty among the modal and non-modal verbs were analyzed by a one-way ANOVA. In addition, t-tests were performed between the non-modal verbs and each of the modals, between *must* and *may*, between *must* and *might*, and between *may* and *might*, to see if a significant difference existed between any of these pairs in terms of their degrees of certainty.

**Results**

The results of ANOVA among the four paragraphs for each modal and non-modal verb showed no significant differences, as shown in Table 2. These results suggest that the degree of certainty that each paragraph was given by the subjects was not associated with the content of the paragraph. Since the only treatment for each paragraph was the difference of modal verbs and the nonexistence of
Atsuko K. Yamazaki

modals, the degree of certainty rated in the paragraphs can thus be regarded as the degree of certainty
given to each modal or non-modal verbs.

Table 2. The results of the one-way ANOVA among the four paragraphs for
each modal and non-modal verbs, and the mean score of each
variable.

<table>
<thead>
<tr>
<th></th>
<th>non-modal</th>
<th>must</th>
<th>may</th>
<th>might</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean of paragraph1</td>
<td>85.778</td>
<td>77.667</td>
<td>42.333</td>
<td>34.833</td>
</tr>
<tr>
<td>mean of paragraph 2</td>
<td>77.067</td>
<td>70.067</td>
<td>36.800</td>
<td>40.667</td>
</tr>
<tr>
<td>mean of paragraph 3</td>
<td>81.611</td>
<td>75.056</td>
<td>49.333</td>
<td>45.833</td>
</tr>
<tr>
<td>mean of paragraph 4</td>
<td>86.100</td>
<td>77.800</td>
<td>46.200</td>
<td>44.75</td>
</tr>
<tr>
<td>F*</td>
<td>1.476</td>
<td>0.974</td>
<td>2.247</td>
<td>1.701</td>
</tr>
<tr>
<td>p**</td>
<td>0.229</td>
<td>0.410</td>
<td>0.091</td>
<td>0.175</td>
</tr>
<tr>
<td>MS (between formats)</td>
<td>297.941</td>
<td>212.129</td>
<td>478.781</td>
<td>452.705</td>
</tr>
<tr>
<td>MS (within formats)</td>
<td>201.912</td>
<td>217.837</td>
<td>213.099</td>
<td>266.061</td>
</tr>
</tbody>
</table>

* df (between formats) = 3  * df (within formats) = 67  ** α = 0.05

The analysis among the modals and the non-modals showed a very significant difference
(F(3,280) = 138.969, p < 0.001), as given in Table 3. As predicted, the non-modals obtained the
highest degree of certainty (M = 82.291%) and must was rated as the second most certain (M =
75.437%). Might was given the lowest degree of certainty (M = 41.648%) and may was rated as
having a slightly higher degree of certainty than might (M = 44.028%). The result of the t-tests
between may and might showed no significant difference (t = 1.105 , df = 70 , p = 0.27), while a very
significant difference was obtained between the non-modal verbs and each modal, between must and
may, and between must and might in t-tests conducted between these pairs, as given Table 4.

Table 3. The results of the one-way ANOVA among non-modal and modal
verbs and the mean scores of non-modal verbs and each modal verb.

<table>
<thead>
<tr>
<th></th>
<th>F*</th>
<th>Mean values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>138.969</td>
<td></td>
</tr>
<tr>
<td>p**</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>MS (between formats)</td>
<td>32038.110</td>
<td>82.291, 75.437, 44.028, 41.648</td>
</tr>
<tr>
<td>MS (within formats)</td>
<td>230.541</td>
<td></td>
</tr>
</tbody>
</table>

* df (between formats) = 3  * df (within formats) = 280  ** α = 0.05

In addition, 49.3% of the subjects attributed higher certainty to might than to may, and 69%
indicated that a sentence without modal verbs expresses more certainty than one with must. More
than 94% of the subjects attributed higher certainty to must than to may, and 90.1% gave higher
certainty to must than to might.
The Pragmatic Function of Modal Verbs in Scientific Papers

Discussion

To a certain extent, the hypothesized order of certainty that each modal verb carries was supported by the results. According to the results expressed in percentages, the verbs can be ordered from the highest degree of certainty to the lowest as the non-modal verbs, *must*, *may*, and *might*. However, the statistical analysis rejected the hypothesized hierarchy for the degree of certainties of the modals and the non-modal verbs. The hierarchical distances in the certainty scale among the modal and non-modal verbs were shown to be different from the hypothesized ones. The certainties that *may* and *might* carry were not clearly distinguishable, and they can be regarded as expressing almost the same degree of certainty. Also, the results demonstrated that *must* had been attributed less certainty and the gap between *must* and *may* was smaller than the hypothesized ones. A diagram of the hierarchy based on the results is shown in Fig. 2.

<table>
<thead>
<tr>
<th></th>
<th>t*</th>
<th>p**</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-modal vs must</td>
<td>4.603</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>must vs may</td>
<td>13.847</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>may vs might</td>
<td>1.105</td>
<td>0.270</td>
</tr>
<tr>
<td>non-modal vs may</td>
<td>18.037</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>non-modal vs might</td>
<td>14.807</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>must vs might</td>
<td>12.451</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Fig 2. A diagram of the certainty hierarchy based on the results for the non-modal and modal verbs.

The significance of this study can be found in the results which show there is no statistical difference between *may* and *might*, and nearly 50% of the subjects attributed higher certainty to *might* than *may*. Many English grammar books, including those for non-native speakers of English, suggest that *might* expresses less certainty than *may*. In particular, many books for English writing and reading for non-native speakers (Yamamoto & Fukutake 1995; Hyodo 1993; The JACET Committee on Teaching Materials 1996) mention that *might* expresses weaker conjecture in comparison with
Atsuko K. Yamazaki

may. Also, they often note that must is used to express inevitability or strong certainty, without giving any clear comparison of must with non-modals (Harada 1994; The JACET Committee on Teaching Materials 1996). This kind of description of must tends to give an impression that a sentence with must expresses higher certainty than one without any modals. However, the results in this study suggest that scientists who are native speakers of English attribute more certainty to a sentence without modal verbs than to one containing must, as described in Halliday (1985).

However, these books are likely to be used as references by non-native scientists for reading and writing scientific papers. If a non-native scientist uses a modal verb in his/her scientific writing to express his/her inference in the way that he/she has learned from non-sociopragmatic aspects described in the above books, there could exist a discrepancy between the certainty that he/she wanted to show and the certainty that the reader infers from the modal verb. The results of this study, therefore, suggest that non-native scientists need to be aware of the pragmatic functions of modal verbs, and grammar books for scientific writing in English should clearly present such aspects of the modals.

Hinkel (1995) points out the difficulty that non-native speakers have in grasping the concepts carried by the use of modal verbs in her research. She mentions the importance of teaching the use of modal verbs by their pragmatic functions, rather than only by their grammatical contexts. Many studies, including Hinkel’s, note that modal verbs can reflect non-native speakers’ cultural and language backgrounds so that they tend to use modal verbs in contexts different from ones in which native English speakers use (Cook 1978; DeGarrico 1986; Gibbs 1990). For a non-native learner, acquiring the pragmatic notion of a modal verb takes time since it reflects the conceptual structures of the language he/she is learning, and the structures often depend on connotations in the society or community where the language is spoken. The acquiring process can be more difficult when the non-native speaker is learning English in his/her first language environment, where he/she is not exposed to enough second-language experiences to deductively learn the meanings of modals easily.

As noted in Halliday (1985), the semantic distinctions among modals involve details and cognizance in various situations. This delicacy associated with the epistemic use of modals hinders the learner from grasping the function of the modals. By the same token, it is not possible to give a deterministic evaluation to each modal since its epistemic usage can reflect complex interpersonal meaning in scientific discourse. However, if the pragmatic meanings of modal verbs can be presented in a way that the learner can intuitively interpret, such as the percentage scales presented in Celce-Murcia and Larsen-Freeman’s book (1983) and this study, the functions of modal verbs can be understood more easily by the learner. In order to establish a clear representation of the pragmatic meanings, studies to quantify the functions of modals used among native English speakers, such as this study, should be conducted more often.

The results of this study also suggest that the teachers should be aware of the uses of modals in their pragmatic frameworks. The meaning and usage of modals are important to the learner who intends to use English in a particular community, but often teachers are not familiar with such
sociopragmatic contexts. For example, it would be difficult for the teacher to judge if the results of this study also represent the standard usage of the modals among engineers, since his/her sociolinguistic norms tend to be associated with the linguistic community consisting of teachers and linguists, but not of engineers. Therefore, a teacher should not address the functions of modals only from the norms of his/her community, but should also pay attention to the sociolinguistic norm of the community where the learner intends to use English.

It is worthwhile to conduct the same kind of research with non-native speakers to see how much certainty they attribute to the modal verbs examined in this study. The author believes that the comparison of results from this study and those from a survey with non-native subjects will give useful insights into the teaching of modal verbs from their pragmatic functions. In particular, the results will help quantitatively identify problems that learners tend to experience in the use of modal verbs. When the problems are quantitatively defined, it will be much easier to deal with them, and the teaching of modal verbs can be more effective. Also, the comparison will give more opportunities for the teachers to be aware of how students use the modal verbs in their pragmatic contexts. The author would like to note that she is in the process of expanding this study in such a direction.

References


**Appendix I: An example of the questionnaire Format I**

Please read the following brief reports on chemical research. Then, indicate how much you think the author of each report is certain of his/her conclusion (the last sentence) by placing an “X” on the percentage scale.

. Several 2,5-disubstituted furans which are known to react with peroxy acids, singlet O, and other active forms of O were tested as potential inhibitors, co-oxidants, or substrates for soybeans lipoxygenase. The furan, 10,13-epoxyocta-deca-10,12-dienoic acid Me ester (I) was converted by lipoxygenase or singlet O or peroxy acid to the acyclic product, Me 10,13-dioxooctadec-11-enoate. Apparently, I is able to interact with an active site of lipoxygenase.
The Pragmatic Function of Modal Verbs in Scientific Papers

The author of this research report is certain of his/her last statement.

\[ \text{Compared to Na-NH}_3 \text{ solutions of similar electron concentration spectra of NH}_3 \text{ solutions} > 10^{-3} \text{ M of Sm and Tb were narrower and broader, respectively. However, NH}_3 \text{ solutions of Tb, Tm and Er of concentration } \sim 10^{-4} \text{M possessed the same band shapes as corresponding Na-NH}_3 \text{ solutions. In addition, the Sm cation spectrum showed Sm to be divalent in these solutions. According to the position and magnitude of one band attributed to the Tb cation, Tb is also divalent in these solutions.} \]

The author of this research report is certain of his/her last statement.

\[ \text{Caerulein was injected into mice in order to study its effect on DNA synthesis activity in the gallbladder epithelium using histoautoradiography, after the injection of labeled thymidine. Higher labeling indexes were observed at 8, 12 and 24 hours after caerulein injection. Apart from its cholecystokinetic effects, caerulein must have exerted a trophic effect on the gallbladder.} \]

The author of this research report is certain of his/her last statement.

\[ \text{The AlCl}_3 \text{ catalyzed cyclizations of benzylaminopropyl bromides were reexamined. Contrary to earlier reports these reactions gave as their major products the corresponding 4-methyltetrahydroisoquinolines rather than the tetrahydro-2-benzazepines. The formation of 4-methyltetrahydroisoquinolines may be due to isomerization of the initially formed carbonium ion prior to ring closure.} \]
Atsuko K. Yamazaki

The author of this research report is certain of his/her last statement.

The dependence of the oxidation rate of Pb films on the pressure of O was investigated by ellipsometric and resistance techniques for O pressures between 200 Pa and 100kPa, and in the temperature range 323--423K. During the first few minutes of the reaction an unusual trend was noted in the observed kinetics. This phenomenon might be due to absorption of O by the metal close to the metal oxide interface.

The author of this research report is certain of his/her last statement.

The electrolytic epitaxial nucleation, growth and coalescence of Cu on (111)-Ag single crystal films was studied. The Cu crystallites are in parallel epitaxial orientation: (111)Cu/ /(111)Ag and (110)Cu/ /(110)Ag. Double positioned Cu crystallites were observed. Crystallites with (100)-orientation were also observed. This phenomenon was probably due to the presence of double positioning boundaries in the Ag films.

The author of this research report is certain of his/her last statement.
The Pragmatic Function of Modal Verbs in Scientific Papers

Appendix II: The numbers and pages of the original abstracts in Chemical Abstracts, vol. 92, 1980.

<table>
<thead>
<tr>
<th>Paragraph</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>210885t</td>
<td>221766f</td>
<td>209281m</td>
<td>215239v</td>
<td>221603d</td>
<td>223166r</td>
</tr>
<tr>
<td>(page)</td>
<td>(213)</td>
<td>(406)</td>
<td>(107)</td>
<td>(606)</td>
<td>(394)</td>
<td>(510)</td>
</tr>
</tbody>
</table>