

The frequency of *we* across academic disciplines and registers in MICASE¹

Mackenzie Chapman, University of California at Santa Barbara
Stefanie Wulff, University of North Texas

Introduction

The study of the use of first and second pronouns within academic discourse has focused primarily on their reference and discourse function. Previous studies by Rounds (1985, 1987a, 1987b) analyzed instructors' use of *we* compared to other personal pronouns in mathematics classes at the University of Michigan, proposing multiple discourse functions and demonstrating a high frequency of *we* among instructors and successful teaching assistants. Using MICASE (Michigan Corpus of Academic Spoken English) to expand upon Rounds' studies, Fortanet (2004) then examined the use of personal pronouns in the corpus as a whole, as well as in a limited corpus containing only mathematics-related speech events to compare her findings with Rounds' study of mathematics lectures. Fortanet found that the use of *we* was half that of other pronouns such as 'I' and 'you' throughout the corpora, proposing that compared to Rounds' results, which displayed a strikingly high frequency of 'we,' this difference may be attributed to the growing acceptance of the use of 'I' in academic language.

These studies raise the question whether the frequency and use of *we* depends at all on the academic discipline and register in which it occurs. In this kibbitzer, we took a first step towards addressing this question by compiling selective corpus samples of three physical science disciplines and three different registers from MICASE.

Methodology

We compiled six corpora from fourteen transcripts selected from the 152 in MICASE, differentiated by one of three physical science disciplines, or one of three registers. We focused on physical science-related speech events in order to be able to compare our numbers to those of Rounds' and Fortanet's investigations.

A multitude of disciplines are provided within the physical sciences, designated by the University of Michigan as astronomy, chemistry, computer science, engineering, geology, mathematics, physics, statistics, and technical communication. After a search of MICASE for speech events from these disciplines, the study focused on engineering, chemistry, and physics simply because there was at least one transcript across three types of speech event detailed below. Thus, separate corpora were compiled for engineering (62,550 words total), chemistry (43,271 words total), and physics (36,835 words total), consisting of six, five, and three transcripts, respectively.

The following three registers were identified:

- a Lecture Corpus containing eight lectures (65,721 words total);
- a Classroom Corpus including three discussion section and lab section events (26,919 words total); and
- a Non-Classroom Corpus with a total of three study groups and office hours (50,016 words total).

¹ We thank John Swales for comments on an earlier version of this kibbitzer, and Elizabeth Buchanan for her kind support with editing this manuscript.

By limiting each corpus to these types of speech events, we aimed to preserve similarities of register within the corpora while providing enough of a difference in registers among the corpora.

Results

Several searches using the concordance program *AntConc* produced the results in Table 1. For comparison, Fortanet's results for the frequency of *we* in MICASE and in the reduced corpus of mathematics-related speech events were added to the table, as were Fortanet's (2004) estimated frequencies for Rounds' results. For a more thorough assessment of the distinctions of the frequencies found, we added the standard deviations calculated for each corpus in this study to the table as well.

Table 1. Occurrences and frequency of *we*.

Corpus	Occurrences of <i>we</i>	Frequency (per 1,000 words)	Standard deviation of <i>we</i> frequency in files within each corpus
Rounds' corpus	907	34	
<i>Fortanet's corpora</i>			
MICASE	7,450	9.7	
Mathematics	452	14.8	
<i>Academic Discipline</i>			
Engineering	923	14.76	10.13
Chemistry	542	12.53	2.25
Physics	341	9.26	0.82
<i>Register</i>			
Lecture	746	11.35	1.95
Classroom	326	12.11	3.27
Non-Classroom	734	14.68	13.88

As demonstrated in Table 1, the frequency of *we* across disciplines remains similar, hovering around approximately 12 occurrences per 1000 words. Upon closer inspection of the Engineering Corpus transcripts, the high standard deviation was likely due to the only non-classroom speech event, a study group transcript containing 33.31 occurrences of *we* per 1,000 words. If the transcript is omitted, the frequency of *we* reduces to 10.67 with a standard deviation of 3.69 for the Engineering Corpus, which would put the frequency of *we* in the same range as the Chemistry and Physics corpora. Within this range of frequencies, Fortanet's results from the mathematics corpus remain slightly higher than these three disciplines at 14.8 occurrences of *we* per 1,000 words. However, as the standard deviation of Fortanet's corpus is unknown, a complete comparison of the corpora remains difficult.

The results for the corpora differentiating register show minimal differences in the frequencies of 'we.' Once again, the high standard deviation of the Non-Classroom Corpus may in part be attributed to the engineering transcript as described above. Once we exclude the above-

mentioned engineering transcript from the Non-Classroom Corpus, the frequency of *we* in the Non-Classroom Corpus comes slightly down to 14.40 occurrences of *we* per 1,000 words, and the standard deviation drops drastically to 0.91. Generally speaking, then, the numbers indicate that *we* is more frequent in the Non-Classroom Corpus compared to the Classroom and Lecture Corpora.

Meanwhile, the Lecture Corpus produced half the number of occurrences of *we* compared to Rounds' corpus, which was also restricted to lectures, and thus places the frequency of *we* found in the Lecture Corpus on the lower end of the range between the Classroom and Non-Classroom corpora. This demonstrates a slight positive trend from lecture to non-classroom registers; though if one considers the range of standard deviation among the results, this trend needs to be interpreted with caution.

Although both Fortanet's findings and the results of this study present evidence to counter the high frequency of *we* found in Rounds' study at an estimated 34 occurrences of *we* per 1,000 words, one cannot ignore the fact that the non-classroom, engineering speech event contained a similarly high frequency of *we* at 33.31 occurrences per 1,000 words. A more in-depth discourse analysis of this study-group transcript would likely provide more insight to determine the significance of this apparent anomaly, as would a more comprehensive study of other non-classroom speech events within MICASE.

Conclusions

The corpora selected from MICASE for this study produced strikingly fewer occurrences of *we* compared to the high frequency of *we* found by Rounds in her transcripts from the 1980s; on the other hand, our results are consistent with those in Fortanet's study. While the differences between the frequencies of *we* in engineering, chemistry, and physics appeared rather insignificant, the slight increase from the Lecture to Non-Classroom corpora may indicate a possible positive correlation of the frequency of *we* with informal registers. However, without an examination of the linguistic contexts of *we* throughout these transcripts and determining their referents and functions, these trends of frequencies between registers cannot be fully explained. Although the differences may indeed be a reflection of the dynamics of the speech events, the distribution of *we* functioning as a representation of a group or in relation to metadiscourse (Fortanet 2004) could possibly differ between registers. Further analysis of the differences in frequency of *we* between registers thus requires more in-depth research of its linguistic contexts and functions.

References

- Fortanet, I. 2004. The use of *we* in university lectures: reference and function. *English for Special Purposes*, 23(1), 45-66.
- Rounds, P. 1985. Talking the mathematics thorough: disciplinary transaction and socio-educational interaction. Unpublished PHD dissertation, The University of Michigan.
- Rounds, P. 1987a. Multifunctional personal pronoun use in educational setting. *English for Specific Purposes*, 6(1), 13-29.
- Rounds, P. 1987b. Characterizing successful classroom discourse for NNS teaching assistant training. *TESOL QUARTERLY*, 21(4), 643-671.