Part 7

Corpus-Based Pedagogical Materials Development:
Guidelines for Teachers and Sample MICASE-Based Materials

Historically, post-secondary ESL/EAP teachers have relied heavily on written discourse and/or their own intuitions about how language functions in academia to create teaching materials for the speaking and listening components of academic discourse. With the availability of a specialized corpus such as MICASE, we now have access to a wealth of authentic examples of the very kinds of interactions we wish to target. Armed with this kind of resource, we can more easily guide our students in learning pragmatically and sociolinguistically *likely* and *appropriate* uses of language, rather than just grammatically correct uses. In this set of guidelines, created specifically with both the corpus and the *MICASE Handbook* in mind, we present a variety of suggestions to help teachers think about how to search the corpus for appropriate excerpts, how to incorporate such excerpts into teaching materials, and where in their syllabus their students could benefit from exposure to this authentic speech. Three sets of sample teaching materials are included to give an idea of the range of possibilities for pedagogical applications using this type of corpus.

Readers may notice that traditional concordance-style lexical and grammatical exercises are not prominent in these samples. We draw on a strong tradition at the University of Michigan's English Language Institute of genre- and discourse-based approaches to EAP teaching. The materials we offer here reflect our inclination to combine corpus-based approaches with discourse-based approaches in developing teaching materials. This is not, of course, to imply that concordance-based exercises are not worthwhile; they do indeed have their place in classroom applications of corpora. Our focus, however, has been elsewhere as will become evident in the materials that follow.

GENERAL GUIDELINES

The first step in any materials development process is to decide *what* to teach, based on what students need to learn. Once the target and teaching goal have been identified, the next step is to find relevant excerpts in appropriate transcripts. Different teaching goals will lead to different starting points; sometimes the teaching goal revolves around a particular situation of use and other times around a particular linguistic structure or pragmatic function. Teachers are encouraged to think in terms of (1) the *situation* in which the language of interest takes place or the language that is specific to or tends to occur in a particular situation, (2) the *functional or pragmatic* use of the language (e.g., humor, advice giving, use of visuals; or more global discourse features, such as those that are characteristic of interactivity, for example), or (3) particular *lexico-grammatical* features of the discourse. Here we provide a few starting points.

Situational

1. Target a speech event, context, or setting. Identify a specific type of interaction or event of interest (e.g., study groups, discussion sections), then use the **Speech Event Type Index** (pp. 18–23) to find all such events.
2. Target particular speaker characteristics. To find events that include the speech of, for example, graduate student instructors, senior faculty, female undergraduates, etc., go to the **Primary Speaker Characteristics Index** (pp. 55–66) to find events where such speakers have a prominent role.

Functional-Pragmatic

1. Target a particular pragmatic function. For example, look up "definitions" or "disagreement" or "advice" in the **Pragmatic Features Overview Chart** (pp. 70–81) for transcripts with this type of content. Use the **Speech Event Abstracts** (pp. 101–252) to select transcripts to investigate further.
2. Target certain discourse style characteristics. For example, in linguistic terms, what seems to stimulate interactivity in a class? Or,

how are monologic lectures structured? Look in the **Interactivity Rating Index** (pp. 34–43) to find highly or mostly interactive (HI or MI) or highly monologic (HM) speech events, then perhaps compare and contrast these events with others in the same speech event category, using the **Speech Event Type Index** (pp. 18–23) as a reference.

3. Contrast functional-pragmatic language across fields or disciplines. For example, use the **Pragmatic Features Overview Chart** (pp. 70–81) and/or the **Keyword Index** (pp. 253–265) to identify transcripts with particular features (e.g., question-and-answer periods) or conversation topics (e.g., experiments) to compare discourse structure, register, vocabulary, etc.

Lexico-Grammatical

1. Target the use of a particular word or phrase, for example an idiom or formulaic expression (e.g., "on the other hand," "by the way"), or a set of related words (e.g., sentential adverbs such as "actually" and "basically"). First, the **MICASE Online** website (or other concordancing program) can be used to search for the word or phrase. (See pp. 16–17, 273, 276–278 for references to using MICASE Online.) How is it used? What collocates with that term? In what kinds of situations is it used—formal or informal? And by whom—undergraduates or faculty? Then from the hits returned, select one or more transcripts and go to that event's **Speech Event Abstract** page to find out more about the speech event and speakers.

2. Target the lexicon of a particular field of study or specialty within that field. Use the **Academic Division Index** (pp. 24–33) to identify all of the events within the field (e.g., biology) or use the **Keyword Index** to further limit the scope of the search to a specialty or subtopic (e.g., genetics; American culture). Use the data to develop a specialized vocabulary lesson that includes terminology unique to that topic that is appropriately contextualized.

3. Focus on a grammatical structure. For example, teachers may want to highlight constructions involving modal verbs, especially as they are used in hedging an opinion or softening criticism. After identifying some modal verbs (*could, should, might,* etc.), use the **MICASE Online** interface (or other concordancing

Materials Development

program) to search for utterances in which these verbs appear. If desired, select only particular speaker and speech event attributes, or display and/or sort according to specified speaker attributes. The **Pragmatic Features Overview Chart** (pp. 70–81) can be used in conjunction with these results to give insight into the transcripts in which these modals are more likely to be used in discussing or evaluating.

In addition to the three broad categories above, this handbook has the added advantage of being able to guide teachers in other directions. At times teachers may want to target certain broad content matter or specific topics of conversation appropriate for their students regardless of the particular language feature they are focusing on. For example, a teacher of students in health professions could limit the search to the biological and health sciences—fields that their students are more likely to be interested in—whether their goal is to teach lecture discourse, politeness features, or certain vocabulary items. To locate discourse in a particular field, refer to the **Academic Division Index** (for general academic areas), or go to the **Keyword Index** and look under the general topic or a more specific subtopic (e.g., immunology or cancer).

Another approach is to simply look to the data for inspiration. Browse the **Keyword Index** for non-content topics of conversation or other classroom situations or events. Some of the categories indexed that may be helpful in this regard include problem solving, equipment failures, and small talk. Browse the **Abstracts** for ideas that can augment an existing battery of resources and teaching materials. Reading through some of the **Abstracts** in the *Handbook* may inspire any number of investigations: examine the ways in which an instructor manages the classroom; observe how students interact with one other as compared to the instructor in discussion classes or small groups; or contrast the informal language of lab sessions with that of more formal settings such as large lectures, focusing on broad characteristics such as discourse structure or more specific features such as vocabulary or grammatical structures.

Thus far, the emphasis has been on using this handbook to find appropriate MICASE data to draw on when creating corpus-based teaching materials. Next we will focus on some of the ways teachers can take that data and actually use it to develop classroom lessons.

DEVELOPING A CLASSROOM LESSON

Here, we explore the process of using the *MICASE Handbook* in conjunction with the transcripts in the MICASE corpus to develop a classroom lesson. Some examples are taken from our own experiences as we created a lesson on opinions; the resulting transcript excerpt, student worksheet, and teacher commentary are included in the sample teaching materials found on pp. 285–318.

An obvious first step in planning for any lesson is to consider the needs of the students. For example, one of the greatest challenges students in academic contexts face is participating fully in discussions, so a lesson on opinions, agreement, and disagreement targets some crucial academic speaking skills. Then consider the types of situations in which exchanges of opinion occur. There are discussions, of course, but discussions can occur in many types of speech events: colloquia, discussion sections, dissertation defenses, and meetings, to name just a few. So it's best not to restrict parameters to only discussion sections. Neither should the search be limited to a particular phrase or construction at the outset. For example, *in my opinion* turns out not to be a very common way of inserting one's opinion. Again, the search should not be based primarily on intuition, but rather left open to all possibilities until sufficient empirical evidence emerges.

SELECTING AND PREPARING EXCERPTS

The next step is to choose one or more speech events from which to extract passages or examples. We encourage the use of a combination of resources in this handbook to find appropriate transcripts:

- Use the **Pragmatic Features Overview Chart** (pp. 70–81) to find events that include the target feature (e.g., advice, discussion). Then look at the **Abstracts** of some of these events to find out more about each event. For example, looking up advice in the chart on p. 70, the first two speech events with numerous examples of advice are Academic Advising and American Culture Advising. The Academic Advising abstract can be found on p. 198.

- Check the **Interactivity Rating Index** to locate events with a desired level of interactivity. For example, in creating the opinions lesson (included here in the sample teaching materials on p. 285), we assumed the more interactive events would have a greater variety of speech and speakers, and a greater likelihood of providing good examples of opinions, and agreement and disagreement.
- Alternatively, any of the other ordered lists of transcripts can also be used as starting points in conjunction with the abstracts to help confirm or narrow the search.

Once one or more speech events that look promising have been identified, go to the abstract, where a synopsis of the event can be found. The abstract describes the content and context of the event, and the summary table at the top of the page summarizes information about the speaker(s) and participants, the recording length, the level of interactivity, and salient or interesting pragmatic features. Note that the abstracts offer a wealth of pertinent and discriminating information that is valuable during various phases of the initial search and throughout the entire development process. For example, from the Academic Advising abstract we learn that multiple students request advice on a variety of topics, including choosing majors and courses, handling a bad grade, and career planning. In addition to advice, other pragmatic highlights listed in the summary at the top of the abstract include examples, questions, and requests.

The value of referring to the abstracts after consulting the Pragmatic Features Overview Chart was also evident during the development of the lesson on opinions. Although "opinions" is not one of the features listed in the chart, opinions are commonly found in discussions, and "discussion" is indexed. Therefore we chose to look at abstracts from several speech events listed as having at least some discussion. The abstract of one event in particular—the Behavior Theory Management Lecture— revealed several appealing features. The professor asks students for their opinions on the effectiveness of grades as a motivator in school and draws an analogy of grades in school to money in the workplace. This topic is accessible and relevant to students and thus would make an appropriate and interesting discussion topic in a variety of classes. In addition, the abstract notes that the students break into small discussion groups at one point and share their conclusions with the class at the end. Another interesting feature of this class is specified in the abstract's sum-

mary table: There are 60 students and 50 of them speak, making for variety—or notable similarity—in student linguistic expression.

The next step is to go to the transcript itself, which can be accessed either via the MICASE Online website (*http://micase.umdl.umich.edu/ m/micase*) or by purchasing a copy of all the corpus texts. (See pp. 16–17 for details on ordering MICASE.) For example, to get to the Behavior Theory Management Lecture transcript from the MICASE Online website, select "Browse MICASE." Under the "Speech Event Type" menu on the left, select "Lecture-Large." Scroll to the bottom of the page and click "Submit Search." Select file LEL185SU066 (Behavior Theory Management Lecture) from the list of large lectures, and then choose "View entire transcript in HTML." From here you can either read the transcript on line or print parts or all of it. Finally, from the chosen transcript, select rich passages that might lend themselves to a variety of lessons, exercises, and activities.

In choosing transcripts and segments to base a lesson on, it is important to think about the appropriateness and accessibility of the subject matter as it relates to both the teaching aims and the students. Criteria might be related to the type of speaker, the number of speakers, the subject matter (not too obtuse or controversial), the linguistic richness of the text (the number of examples present), or the amount of jargon or slang (whether you want a little or a lot).

For example, for the lesson on opinions (included in the sample materials), after a few promising transcripts were identified, the excerpt ultimately chosen satisfied the following criteria:

- The excerpt is relatively short and dense.
- Everyone present participates.
- There are a number of different speakers.
- All speakers give an opinion.
- The topic is non-technical and accessible, with little or no jargon or slang.
- All speakers are students.

On the other hand, for certain purposes, teachers may want to target discussions that are moderated by a graduate student instructor or professor in order to analyze, for example, how they manage the discussion or how students interact with them in that context.

Materials Development

The students may also present another set of criteria or constraints. Bear in mind not only the students' language proficiency level, but also their academic interests, maturity level, and cultural background when selecting a transcript or excerpt. Students in the humanities and "soft" sciences may find excerpts from some "hard" science classes too technical or dense. Conversely, for a class comprised entirely of engineers, it may be tempting—and possibly advantageous—to choose a transcript from an engineering class. However, some MICASE users argue that it is best to utilize content that is unfamiliar to the students, or at least not within their specialty, because students can be distracted by the factual correctness of the content at the expense of the particular language points in focus.

In any case, teachers should choose a transcript and passage to their liking but are encouraged to keep in mind others that have surfaced during the search because additional excerpts may inform the teacher's own conclusions about the target features and may be helpful in developing additional exercises. Beware, however, that using excerpts from transcripts other than the one that forms the main basis of the lesson or taking statements out of context may make it more difficult for students to follow the excerpts.

After settling on a transcript, locate some passages that are especially rich. We suggest printing out some portions of the transcript and then making notes and highlighting phrases or utterances. For example, again using the opinions lesson as a model, after highlighting all the expressions of opinion that could be identified, it immediately became obvious that the overwhelming majority of these included some form of *I think*. (See pp. 289–295 for the edited transcript exerpt.) Note contrasts (e.g., strong vs. weak opinions) and other features of interest (e.g., expressions of agreement and disagreement); mark strategic language (e.g., strategies for avoiding giving an opinion) and other linguistic devices (e.g., questions, used both for giving and for soliciting opinions). We also suggest noting grammar points or other interesting features and looking for trends, useful phrases or idioms, etc. In the course of reading and analyzing the transcript, noteworthy features reveal themselves, so teachers should try not to set out with a narrow idea of the exact features they are interested in, but rather let the data speak.

The final step is to use the chosen excerpts, along with the notes, to develop discussion questions and exercises, possibly in the form of a

worksheet that is attached to a copy of the text. In the section that follows, we provide a list of suggested exercises and activity types.

SUGGESTED EXERCISES, ACTIVITIES, AND HOMEWORK ASSIGNMENTS

The following exercises and activities provide students with opportunities to practice and apply their new skills in the classroom.

Written Exercises

These are generally written exercises, suitable for inclusion in a worksheet, that can be done individually or in pairs/small groups.

- *Identification:* Instruct students to circle, underline, or otherwise mark words, phrases, or passages in the transcript (e.g., grammatical constructions, problem sounds).
- *List:* In conjunction with identification, have students make a list of all of the phrases they can find in the text that perform a particular function or fall into a stipulated category.
- *Cloze or other fill-in-the-blank exercises:* Delete words or phrases from excerpts from the transcript and have students fill in the blanks.
- *Ranking:* Have students rank a series of items according to certain criteria, such as formality, strength, etc.
- *Multiple choice questions.*
- *Writing:* Have students create or revise sentences, dialogues, or entire scripts to incorporate features of interest.

(*Note:* Several of these kinds of exercises are included in the worksheet for the Opinions Lesson—identifying, listing, ranking, and writing.)

Speaking Activities: Speaking Practice in Communicative Contexts

- *Group discussions:* Assign topics for small or large groups of students to brainstorm and/or discuss some aspect of the target.
- *Script reading:* Use the transcript as a script; assign or have students choose roles and then act out the scene.

Materials Development

- *Script reading with a twist:* Change some characteristic(s) of one of the speakers and have students act out the transcript, with improvisation.
- *Predict and role-play what will happen next in the transcript.*
- *Role play:* Have students identify some situations where they will have to use the skills learned in this lesson and act them out.
- *Prediction or history:* Show the transcript one part at a time and have students predict what will happen next or speculate on what may have happened just before.
- *Surveys:* Have students do a survey or take a poll of their classmates about some aspect of the topic.
- *Interviews:* Have students practice for a homework assignment that requires them to do an interview relating to the excerpt with a native speaker or classmate or more advanced English learner.

Computer-Based Activities

A corpus is a computerized resource, and since MICASE is also freely available online, there are no restrictions to using the corpus in the classroom, other than the availability of a computer classroom or lab. If a computer classroom is not available for teaching, then the following activities can be assigned to students to complete on their own time, provided they have access to the Internet. Alternatively, teachers can also adapt the following exercises by conducting the searches of the corpus themselves and then preparing printouts of selected sets of results or worksheets based on those results and findings.

Most of the references in this section refer to the online search interface for MICASE, available at *http://micase.umdl.umich.edu/m/micase*; however, readers should be aware that while this search interface is sophisticated in the number of speaker and speech event parameters that can be included in a search, as well as some other unique functions, there are a number of limitations to the online interface in terms of the types of searches and subsequent data manipulation that can be done. For example, this search interface does not support Boolean searches or wild card searching, and result sets larger than 500 hits cannot be sorted. For these reasons, many users may prefer to purchase an off-line version of the corpus texts and use them with a desktop concordance program such as WordSmith Tools or MonoConc Pro.

What follows is just a sampling of some hands-on activities that are intended to spark ideas for further activities and adaptations.

- Ask students to look up two or more closely related words to explore the differences in usage: e.g., *however* and *but*; *although* and *even though*; *nevertheless* and *nonetheless*; *on the one hand* and *on the other hand*; *likewise* and *similarly*; *instead* and *rather*; *important, vital, critical, crucial*. Once they get results, experiment with sorting the results by one word to the left or one word to the right to see which contexts are more revealing of usage patterns.
- Find examples and compare usage of the oft misused verbs *say*, *talk*, and *tell* and have students discover the patterns and create their own exercises to practice using these verbs. (Or, create a worksheet using data from the corpus.)
- Ask students to look up a slang word or other word of interest to them—e.g., *cool*, *awesome*, *totally*, *sucks*, *stuff*—and find out (a) which speech events it occurs in and relative frequencies (using "browse" mode to find all transcripts with that item and number of times it occurs); (b) which speakers are using it; and (c) any interesting observations about its semantic, pragmatic, or grammatical context.
- Look up a grammatical frame such as *so* +ADJ *that . . .* or *such +* NP *that*, and find out what lexical items most commonly occur in those slots. The easiest way to do this is to use the proximity search: the main search term would be *so* or *such,* the context word or phrase would be *that,* within two or three words to the right.
- Ask students to pick a common word they hear around them but aren't exactly sure how to use themselves and then look that word up in MICASE and try to come up with their own hypotheses about its meaning and function. Some possible candidates that come to mind for this type of exercise are: *actually* or *anyway*.
- For any lexical item or phrase with enough examples in the corpus, another possible step is to compare the use of that word or phrase in different fields or academic divisions or across different categories of speakers. If the search yields more than 500 results and you want to resort them, try doing the search again but limiting some of the speech event parameters to end up with a more

manageable result set (e.g., only choose transcripts from the humanities; or only look for lectures or office hours, etc.).

- With the online version of MICASE, one of the useful features is the ability to go from the "Key word in context" (KWIC) line directly to the speaker's full turn (by clicking on the line number from the initial results list), and from there it is easy to get to the entire transcript if more context is desired.

Homework Assignments: Opportunities for Students to Apply Their New Skills Outside the Classroom

- Ask students to come up with a few topics that they are really interested in learning more about and then go into the community (if ESL) or to others in the class or more advanced learners (if EFL) to get information about the topic, perhaps in a survey or interview format. (See the homework assignment in the Opinions Lesson.)
- Ask students to observe and take notes of examples of the target interaction in the community (ESL) or in movies or TV programs (EFL).
- Ask students to find other examples of similar interactions either in MICASE or in their own lives.
- Ask students to use MICASE Online if they have further questions about an aspect or feature of this or other interactions.
- Ask students to use MICASE Online to explore the meaning and use of a specific word or phrase.

PREPARING THE TEXT (TRANSCRIPT EXCERPTS)

Teachers will probably want to provide a copy of portions of the transcript (and perhaps a companion worksheet) to students at some point during the lesson, and it is likely that it will be used repeatedly. We suggest editing or "cleaning up" the transcript so that it is easier for students to read, but keep in mind the value of the transcript's authenticity. In addition, the source of the transcript should be included in the form of a citation, along with a notation if it was "adapted from" the original or "revised" in some way; this not only credits the proper sources

but also validates the transcript's authenticity for the students. The recommended citation for MICASE is: R. C. Simpson, S. L. Briggs, J. Ovens, and J. M. Swales. (2002). *The Michigan Corpus of Academic Spoken English*. Ann Arbor: The Regents of the University of Michigan. The name of the transcript and file may also be included, e.g., Behavior Theory Management Lecture (LEL185SU066).

Revisions we have found helpful:

- Number the lines or turns.
- Capitalize (i's, and first word in each turn, etc.).
- Change or reformat some transcription conventions (overlapped speech, laughter, pauses, etc.).
- Adjust the format (line spacing, indents, including hanging indents for each speaker, etc.).
- Assign names to speakers. (The header of the transcript can be checked to determine the speaker's gender, if desired.)
- Delete some dysfluencies, particularly if they are excessive or unduly intrusive (i.e., *uhs* and *ums,* false starts, repetitions, etc.).
- Correct the grammar. Even though the beauty of MICASE is its authenticity, sometimes it just seems wrong to include a grammatically incorrect utterance in teaching materials. Consider the pros and cons.

There is some debate among corpus linguists and language teachers about how much a transcript should be cleaned up, if at all. Should any or all of the dysfluencies be deleted? Although these are natural and some would argue important features of spoken discourse, they can be distracting to students when reading a transcript. What about grammatical mistakes or non-standard elements in the transcript? Can teachers present sample utterances but teach students that they are incorrect, and thereby, in essence, model bad grammar? In making these decisions, one must balance the desire to show real English as it is, imperfectly spoken (which can be comforting to learners), with the desire to teach correct grammar. In fact, this makes a good point for discussion with students. What is their opinion? For an example of how this issue was handled in the Opinions Lesson included in the sample materials, the reader may want to refer to MICASE Online to compare the original transcript with our cleaned-up version.

Materials Development