

TR-IT-0081
A Bilingual Set
of Communicative Act Labels
for Spontaneous Dialogues

Mark Seligman, Laurel Fais, Mutsuko Tomokiyo

November, 1994

Abstract

This document presents a tentative set of Communicative Acts (CAs) which has been used to label spontaneous dialogues in both English and Japanese. A Communicative Act is a communicative goal or aim which (according to native judgments) can be expressed in language L by a distinctive set of conventional cue patterns in specified discourse contexts. Communicative Acts are thus similar to speech acts, and similar to the pragmatic categories often called IFTs (illocutionary force types) at ATR. However, we restrict our attention to communicative goals which can be explicitly expressed via conventional surface cue patterns, thus excluding goals which are expressed using one-time-only combinations, goals which are expressed only implicitly, or goals which can only be defined in terms of relations between utterances. We describe methods of discovering and revising CAs which depend on native judgments concerning essential equivalence of meanings and functions of cue patterns in context. While these judgments are subjective, they concern shared conventions regarding objectively observable objects – the cue patterns and contexts. Thus a consensus can be expected to emerge during repeated revision. In this important respect, the methodology is data-driven or corpus-based. The present study also emphasizes comparison of CAs in English and Japanese. We find that most of our proposed CAs are valid for both English and Japanese: only two out of 27 CAs seem to be monolingual for our corpus. We begin by introducing CAs and briefly describing the background, goals, and status of our research. In later sections, we discuss our methodology in greater depth; we list our current Communicative Acts, with descriptive glosses, representative sets of surface patterns, examples, and other information; we present two labeled dialogues in English and two in Japanese; and finally, we provide an Appendix describing work in progress.

ATR 音声翻訳通信研究所
ATR Interpreting Telecommunications Research Laboratories
©ATR 音声翻訳通信研究所 1994
©1994 by ATR Interpreting Telecommunications Research Laboratories

Abstract

This document presents a tentative set of Communicative Acts (CAs) which has been used to label spontaneous dialogues in both English and Japanese. A Communicative Act is a communicative goal or aim which (according to native judgments) can be expressed in language L by a distinctive set of conventional cue patterns in specified discourse contexts. Communicative Acts are thus similar to speech acts, and similar to the pragmatic categories often called IFTs (illocutionary force types) at ATR. However, we restrict our attention to communicative goals which can be explicitly expressed via conventional surface cue patterns, thus excluding goals which are expressed using one-time-only combinations, goals which are expressed only implicitly, or goals which can only be defined in terms of relations between utterances. We describe methods of discovering and revising CAs which depend on native judgments concerning essential equivalence of meanings and functions of cue patterns in context. While these judgments are subjective, they concern shared conventions regarding objectively observable objects – the cue patterns and contexts. Thus a consensus can be expected to emerge during repeated revision. In this important respect, the methodology is data-driven or corpus-based. The present study also emphasizes comparison of CAs in English and Japanese. We find that most of our proposed CAs are valid for both English and Japanese: only two out of 27 CAs seem to be monolingual for our corpus. We begin by introducing CAs and briefly describing the background, goals, and status of our research. In later sections, we discuss our methodology in greater depth; we list our current Communicative Acts, with descriptive glosses, representative sets of surface patterns, examples, and other information; we present two labeled dialogues in English and two in Japanese; and finally, we provide an Appendix describing work in progress.

1 Introduction

This document presents a tentative set of Communicative Acts (CAs) which has been used to label spontaneous dialogues in both English and Japanese.

We begin by introducing CAs and briefly describing the background, goals, and status of our research.

1.1 What is a Communicative Act?

We can define a Communicative Act as follows:

- A Communicative Act is a communicative goal or aim which (according to native judgments) can be expressed in language L by a distinctive set of conventional cue patterns in specified discourse contexts.

INFORM, ACTION-REQUEST, and YN-QUESTION are typical Communicative Acts. The communicative goals which they respectively represent are (roughly) (1) “speaker wishes to convey new information to hearer”; (2) “speaker wishes to convey to the hearer that speaker wishes hearer to perform some action”; and (3) “speaker wishes to convey to hearer that speaker wishes hearer to convey to speaker whether a specified proposition is true or not”. According to the shared conventions of English, they can respectively be expressed using (1) declarative clause syntax, sometimes with various sorts of ellipsis; (2) using expressions like “would/will/can/could you [verb phrase, bare infinitive]” or “please [verb phrase, bare infinitive]”; and (3) using inverted interrogative clause syntax and/or special prosody. Essentially the same goals can be expressed by appropriate cue patterns in Japanese.

Communicative Acts are thus similar to speech acts, and similar to the pragmatic categories often called IFTs (illocutionary force types) at ATR. However, we use this new terminology to stress several differences in principle and emphasis.

The most important difference arises because we wish to explore the limits of discourse analysis based on surface clues. To test these limits, we expect that early experiments will avoid the use of programs for plan analysis. Thus in our present study we cannot expect to recognize all communicative goals. Instead, we restrict our attention to only those communicative goals which can be expressed using conventional linguistic cue patterns, that is, fixed cue patterns which can be memorized and used repeatedly as part of speakers’ shared knowledge of a given language. We reserve the term Communicative Act for only such conventionally expressible goals. Communicative goals which cannot be described as Communicative Acts include utterance goals which are expressed non-conventionally (using one-time-only combinations); or goals which are expressed only implicitly; or goals which can only be defined in terms of relations between utterances.

We will describe methods of discovering and revising CAs which depend on native judgments concerning essential equivalence of meanings and functions of cue patterns in context. While these judgments themselves are subjective, they concern shared conventions regarding objectively observable objects – the cue patterns and contexts. Thus a consensus can be expected to emerge during repeated revision. In this important respect, the methodology is data-driven or corpus-based.

A second important difference in emphasis arises from our interest in comparing CAs in English and Japanese (and potentially in other languages). While CAs are defined by monolingual conventions for expressing certain goals using certain cue patterns, it is possible to compare the conventions in language A with those in language B by comparing the respective goals and cue patterns. In

practice, we find that most of our proposed CAs are valid for both English and Japanese: only two out of 27 CAs seem to be monolingual for our corpus.

We discuss all of these points more fully below. For the moment, a few examples will be helpful. Here is an English utterance labeled with CAs:

- A : so you'll wanna take the subway north to Sanjo station (INSTRUCT) and (you can catch) you can catch the [ah] subway right there (SUGGEST)

And here are the cue patterns which were used to identify each CA:

- INSTRUCT: "you will want to [VP]"
- SUGGEST: "you could/can [VP]"

Similarly, here is an example for Japanese:

- え (EXPRESSIVE) すいません (APOLOGY) もう一回お願いします (ACTION-REQUEST)

And here are the corresponding cue patterns:

- EXPRESSIVE: "え"
- APOLOGY: "すいません"
- ACTION-REQUEST: "[conjunction/adv] お願いします"

1.2 Background and Related Research

In evolving the methodology explained below and developing our current CA set, we of course referred to related work at ATR and in the literature.

Three ATR-based studies were especially helpful starting points: [Kume and Sato 1989], [Sadanobu et al. 1991], and [Nagata et al. 1993]. [Fais 1991] and Toshiaki Iwadera both derived modified "IFT" (illocutionary force type) listings from these studies which became immediate sources for our work. (Fais adapted material from the first and second studies, while Iwadera used labels from the first and third [personal communication].)

Tomokiyo independently compiled a list of "Discourse Labels" from various sources [Tomokiyo and Morimoto 1992, Tomokiyo 1993a, 1993b] which became a third immediate source.

We began the present study by comparing the Fais, Iwadera, and Tomokiyo lists. (The comparison records are available from the authors.) We also reviewed and incorporated ideas from [Iida and Arita 1992], Y. Sobashima [personal communication] and [Myers 1990] (who in turn cites [Searle 1975] and [Wierzbicka 1987]). As the study continued and the methodology evolved through repeated attempts to apply various label sets to our bilingual corpus, we found a great deal of help, both theoretical and specific, in [Stenström 1994] and [Leech 1983]¹.

Our aim has been to find a *principled* way to synthesize these various sources in order to meet the goals outlined below. We stress the need for principled and structured discovery and evaluation, since it was not our intention to simply add an intuitive label set to the several sets available in our sources.

¹Special thanks to Masato Ishizaki for bringing the last two valuable sources to our attention. Ishizaki and Toshiaki Iwadera both contributed considerable time to discussion of labels for the Japanese dialogues. While retaining responsibility for the current analysis, the authors wish to express their appreciation for this contribution.

1.3 Goals

Why do we want to recognize “conventionally expressible communicative goals” in dialogues? We can list at least six major aims. The first three relate to translation, and the last three relate to speech processing.

Concerning translation, we want to:

- **Identify the CAS of the Current Utterance.**

CA analysis of the current utterance is necessary for translation. For instance, in our analysis, the English pattern “can you *?” may express either an ACTION-REQUEST or a YN-QUESTION. Resolution of this ambiguity will be crucial for translation into Japanese.

- **Identify Closely Related Utterances.**

Utterances in dialogues are often closely related: for instance, one utterance may be a prompt and another utterance may be its response; and the proper translation of a response often depends strongly on identification and analysis of its prompt.

For example, Japanese *hai* can be translated as *yes* if it is the response to a YN-QUESTION, but as *all right* if it is the response to a REQUEST.

Further, the syntax of a prompt may become a factor in the final translation. Thus, in a responding utterance *hai, sou desu* (meaning literally “yes, that’s right”), the segment *sou desu* may be most naturally translated as *he can, you will, she does*, etc., depending on the structure and content of the prompting question.

The recognition of such prompt-response relationships will require analysis of typical CA sequences. These are the subject of study in progress (see Appendix). For now, however, we concentrate on description of the CAs themselves as a prerequisite to sequential studies.

- **Analyze Relationships Among Segments and Fragments.**

Early processing of utterances may yield fragments which must later be assembled to form the global interpretation for an utterance. CA sequence analysis should help this assembly, since we hope to learn how CAs typically group together. Again, description of CAs in isolation is a prerequisite for sequential study.

Concerning speech processing, we want to:

- **Predict CAS to aid speech recognition.**

If we can predict the coming CAs, we can partly predict their surface patterns. This prediction can be used to constrain speech recognition. For example, if we can predict the relative probability that the current utterance is a YN-QUESTION as opposed to an INFORM, we may be able to differentiate utterance-final *ka* (a question particle) and utterance-final *ga* (a conjunction or politeness particle), which are often very similar phonetically. Once again, we need to study CAs in isolation before we can effectively study their typical sequences.

- **Provide conventions for prosody recognition.**

Once spontaneous data is labeled, speech recognition researchers can try to recognize prosodic cues to aid in CA recognition and disambiguation. For instance, they can try to distinguish INFORMs and YN-QUESTIONS according to their F0 curves – a distinction which would be especially useful for recognizing YN-QUESTIONS with no morpho-syntactic markings.

- **Provide conventions for speech synthesis.**

Similarly, speech synthesis researchers can try to provide more natural prosody by exploiting CA information. Once relations between prosody and CA has been extracted from

corpora labeled with CA information, they can attempt to supply natural prosody for synthesized utterances according to the specified CA. For instance, YN-QUESTIONS and CONFIRMATION-QUESTIONS (including English tag questions) can be made to sound more natural. (An ongoing project in this area is described in the Appendix below.)

1.4 Status of the Current CA Set

The current CA set is not official in any sense. And it is far from perfect. While it is now relatively stable, having been applied to numerous English and Japanese dialogues, we do not intend it to be final, but fully expect it to evolve during use. Readers should carefully note the date of this report to be sure of getting the latest version.

Further, the current CA set is not intended to be a complete set for either language. Instead, it attempts only to cover the relevant corpora. The dialogues examined to date are from the EMMI-ATR corpus [Loken-Kim et al 1993a, 1993b]. EMMI, the Environment for Multi-Modal Interactions, permits both telephone-only and media-aided two-way communication experiments (see the Appendix). We have examined both sorts of dialogues.

With this orientation complete, we can now go on to discuss our methodology in greater depth. This will be the purpose of Section 2.

In Section 3, we list our current Communicative Acts, with descriptive glosses, representative sets of surface patterns, examples, and other information.

In Section 4, we present two labeled dialogues in English and two in Japanese.

And finally, we provide an Appendix describing work in progress.

2 Defining and Revising Communicative Act Sets

This section explains the principles which guide the definition and repeated revision of CA sets. As mentioned, our approach is distinct from prior attempts to define speech acts in two major respects:

- We focus on communicative goals which are conventionally expressible via surface linguistic cues, ignoring those which are not.
- We hope to compare such goals in English and Japanese, and perhaps other languages.

We now discuss these points in order.

2.1 CAs Are Associated with Conventional Cue Patterns

A *cue* can be any aspect of the surface syntax, morphology, or prosody of Japanese, English, etc. A *cue pattern*, i.e. combination or configuration of one or more cues, *expresses* a Communicative Act. For instance, the morphological cue pattern “[clause] か” (*ka*) can be used to express the YN-QUESTION Communicative Act in Japanese. The cue pattern “could you (please) [VP, bare infinitive]” can express the ACTION-REQUEST CA in English.

To repeat, aspects of prosody can be valid cues. For instance, in both English and Japanese, a certain F0 curve can express the YN-QUESTION CA, even when morpho-syntactic cues are absent. When sufficient non-prosodic distinguishing cues are present, prosodic information can be associated with cue patterns to add naturalness or aid recognition. However, since we do not yet have acceptable notations for prosodic features, we will concentrate on text-based cues in this report.

For early experiments on automatic cue pattern recognition in text transcriptions of spontaneous dialogues (see Appendix concerning Discourse Context Analysis), cue patterns were described very simply as expressions in a regular grammar – that is, as patterns with stars representing wild cards: “would you please*”, etc. This approach is efficient and worked surprisingly well. However, it is clearly insufficient to permit recognition of all the cue patterns which may interest us: it will not suffice to recognize more abstract cue patterns like subject-verb inversion (as required for recognition of the English CA YN-QUESTION), etc. We anticipate that future CA recognition programs will have access to, and the ability to match against, the full output of analysis programs, including parse trees, features, etc. The cue patterns presented below use regular grammar or informal descriptions of abstract grammatical features as convenient.

Segmentation information is needed to recognize cue patterns consistently. For instance, we will analyze *Yes, you can* as two CAs, YES plus INFORM, rather than as a single unified response. We assume that the input will be segmented in a preprocessing stage, before assigning CA labels.

2.1.1 Cue-To-CA Mappings are Many-to-many

We may occasionally find cues which express only one CA and CAs which are expressed by only one cue or cue pattern. Much more typically, however, the mapping between cues and CAs is many-to-many. Thus, in the general case, we must consider a set of possible cue patterns for a given CA, or a set of possible CAs for a given cue pattern. Here we focus on cue pattern sets for specific CAs.

For instance, we presently analyze “can you [VP]”, like “could you [VP]”, as a cue pattern which can express a CA ACTION-REQUEST in English. Thus the cue pattern set for ACTION-

REQUEST would contain both patterns, among others. (However, regarding subtle differences between such patterns, e.g. relating to relative politeness, see below.)

2.1.2 Distinct CAs Must Have Distinct Cue Sets

Our CA definition specifies that distinct CAs can be recognized only if their proposed cue-pattern sets are distinct (not identical).

Thus one possibility is that cue sets are disjoint (have no members in common). In this case, no confusion is likely. However, cue sets can intersect, even to a large degree. In this case, cue patterns which are not in both sets become important for distinguishing the relevant CAs.

A special-case way in which cue sets can be distinct involves hyponymy, or degrees of specificity. [Knott and Dale 1992:18] make this point in relation to relational expressions in discourse such as *in short* or *in sum*, but it can apply to CAs as well:

If one phrase can always be substituted for another, but not vice versa, then the latter phrase should be classified in a category subordinate to that of the former phrase. In this way a taxonomy of synonyms and hyponyms can be constructed. [Knott and Dale 1992:18]

The claim is that *in sum* is more specific than *in short*, *in brief* and similar summarizing expressions because it can only be used at the end of a discourse segment. By examining many such relationships, it may be possible to recognize a taxonomy of CAs, in which certain CAs are more general and others are more specific. We briefly discuss this possibility below.

2.1.3 Precise Pragmatic Description is not Crucial

Definition and recognition of a Communicative Act does not depend on precise pragmatic specification of the communicative function in question. If native speakers can agree that one or more cue patterns can conventionally express a certain communicative function in context, and the proposed set of cue patterns is non-identical to all other proposed cue pattern sets, then a distinct CA can be hypothesized. The precise pragmatic description or gloss of the communicative function can be investigated later. While each CA also receives a mnemonic name based on the gloss, the name does not define the CA any more than the gloss does: INSTRUCT, for instance, is the name we now use for the CA associated with giving transportation directions in our English corpus (see below); but the CA would be the same one if it were named DIRECT, or CA23, instead.

When several cue patterns are proposed for a given CA, native speakers must judge that they can all express the communicative goal in question (whatever its precise description may turn out to be) in the relevant contexts. However, it is not necessary to claim that these cue patterns are functionally equivalent in every respect. For instance, as already noted, we expect most English speakers to agree that "could you *" and "can you *" both can express an ACTION-REQUEST; but many may feel that "could you *" is more tentative and thus more polite than "can you *" We are aware of many such differences; some of them are indicated below as parameters for specific CAs. Future CA sets may try to capture some such differences by finely subcategorizing current CAs.

2.1.4 Methodology I

Let us make a first attempt to summarize our CA basic discovery methodology. Several refinements will be added below.

- To discover a CA set for corpus C in language L:
 - Collect conventional cue patterns which appear to express communicative goals;
 - Based on native judgments concerning their appropriateness for expressing communicative goals in context, sort the collected cue patterns into distinctive cue pattern sets. In a specified discourse context, natives should judge that the expressed communicative goal remains essentially the same in paraphrases based upon set members. It is more important to recognize that the expressed goals remain the essentially same than to describe or name them exactly. (Minor variations in the expressed goal, e.g. regarding politeness, can be tolerated if the essential function is judged to remain constant.)
 - Give a name and descriptive gloss to each proposed cue pattern set.

2.1.5 Constraints and Heuristics

The essence of our CA discovery procedure is to collect cue patterns, sort them according to the goals they can express in context, and then try to describe the goals.

But there are many possible cue pattern candidates, so additional constraints and heuristics concerning cue patterns are welcome. Further, heuristics concerning plausible goals can also be helpful, in two roles: they can help to justify or filter descriptive glosses when these are attempted; and they can provide a top-down element to the discovery procedure, a way of working from goals toward cues which can complement the cue-to-goal discovery procedure described above.

Cue Patterns Must Be Conventional An important constraint on cue patterns is included in the CA definition: proposed cue patterns for expressing communicative goals must be *conventional*. That is, they must be configurations of cues which are used *repeatedly* in a corpus to express a given communicative goal, rather than expressions composed one time only according to the productive capacity of the language. Thus they may be atomic cues containing a single element (such as morphemes or lexemes); but if they are composed of several elements they must be relatively fixed.

We assume that the mapping between a cue pattern and a communicative goal is memorized by a speaker, or listed for a program, as an element of linguistic competence. In this sense, cue patterns and their mappings to communicative goals are “fixed” or “idiomatic”; but this description is sometimes deceptive, since some cue patterns appear to be compositionally analyzable, e.g., “I’d like you to tell me *” and “I’d like to ask *”. When we propose such cue patterns as members of a CA set, it is because they appear repeatedly with only minor and specifiable variation. Thus their mapping to the relevant communicative goal can indeed be treated as a frozen fact about the language. Further, for many such cue patterns, their use to express the relevant communicative goal is not predictable: for instance, “I’d like you to *” is acceptable as a request pattern in English, but a direct translation would not be acceptable in Japanese. Similarly, from the Japanese side, while *分かりました* (*wakarimashita*) appears to express “I have understood” according to the standard grammatical rules, learners of Japanese cannot know the conventional use as a standard acknowledgment (comparable to English *I see*) without experience or instruction.² *Performatives* often provide examples of such apparently interpretable but really fixed and pragmatically unpredictable expressions (as when an English-speaking minister says, *I now pronounce you man and wife*).

² Another confusing point about such patterns is that they can simultaneously express several CAs at once. For instance, in our analysis, “I’d like you to tell me *” and “I’d like to ask *” simultaneously expression INFORM CA because of their declarative syntax; ACTION-REQUEST because they use the pattern “I’d like (you) to *”; and WH-QUESTION because they include the more specific cue patterns “I’d like you to tell me *” and “I’d like to ask *”. Regarding multiple interpretations, especially those involving expressions at different levels of specificity, see further below.

Our focus on conventional cue patterns and their mappings to goals has another consequence: one-time-only combinations will *not* be treated as CAs, even if they actually seem to express a relevant communicative goal. To illustrate, we can use the WARNING CA (which happens not to appear in our corpus): A speaker who says *Watch out for the dog!* or 猛犬注意! (*Mouken chuui!*) is using a conventional cue pattern for this CA. However, one who says *A mean-looking dog is right behind you!* or 質の悪い犬が... (*Tachi no warui inu ga ...*) is using only an INFORM CA, even though the deeper force, goal, or purpose of the utterance may certainly be understood in context as a warning (that is, as information about something undesirable which is purposely provided so that the hearer can try to avoid the problem). [Leech 1983: 222-223] reminds us that interpretation of the deep force in such cases depends heavily on subjective judgment: if the hearer is known to love big dogs with sharp teeth, the utterance may not be a warning at all.

A Priori Goal Descriptions as Heuristics Heuristics regarding communicative goals are provided by abstract characterizations of communicative goals in the literature. For the reasons already given, we do not wish such descriptions to be central in our definition of CAs; but they can usefully prompt us to investigate a proposed communicative goal. If that goal is indeed expressible using conventional cue patterns, then a CA can indeed be hypothesized.

We have found the categorization of English “illocutionary verbs” in [Leech 1983: 198-228] to be quite useful in this goal-side heuristic role: we agree with Leech that the verbs used in English to talk *about* speech acts are useful clues to the conventional speech acts themselves, though a perfect correspondence should not be expected. (We have included with each CA description below examples of related illocutionary verbs.)

Leech presents a 5-way categorization of illocutionary verbs. Using it as a reference point, we can try to clarify what “communicative goals” are. Leech’s categories include (1) Assertive and (2) Expressive, reflecting a basic distinction between utterances whose goal is mainly to inform and utterances which have other purposes, such as the expression of emotion. The remaining three categories are (3) Directive (which can be discussed using verbs like command, request, or urge); (4) Rogative (discussed using verbs like ask or inquire); and (5) Commissive (which can be discussed using verbs like offer or promise). The utterances which these verb types describe express (3) the speaker’s more specific desires and goals, especially desire for other people’s actions; (4) the speaker’s desire that the hearer provide information (which may be seen as a special case of desired action); and (5) the speaker’s reaction to the hearer’s goals and desires (the speaker’s willingness to perform actions for or with the hearer, and also perhaps – though Leech doesn’t mention it – to permit actions; suggestions, instructions, and advice about how the hearer may reach a goal also fall into this area).³

Thus, when we describe CAs as “communicative goals” which can be expressed using conventional cue patterns, we focus broadly on high-level utterance goals – To inform, or not to inform? – and further on more specific goals or reactions to goals which an informative utterance may inform about: the speaker’s goals to get other people to provide information or perform other actions, or the speaker’s reaction toward the hearer’s goals, indicated via offers or promises, suggestions or advice or instruction, etc.

2.1.6 Methodology II: Evaluating and Revising CAs

Because native judgments of cue pattern usability in context, rather than exact pragmatic description, is central in our methodology for recognizing CAs, the methodology is, in an important sense, bottom-up or data-driven: while we may use abstract categorizations of communicative

³Leech presents all five categories as sisters, but it is also possible to view the last three as subcategories of the general Assertive type, since they all convey information about goals and reactions to them.

goals heuristically, it is the discovery of cue patterns used consistently in the corpus which ultimately justifies a proposed CA.

How are CAs evaluated? By verifying that the proposed mappings between proposed cue patterns and proposed CAs in the corpus are consistent. Let us now attempt to refine the methodology for CA discovery introduced earlier by specifying the following evaluation procedure for CAs.

- To evaluate a set of CAs:
 - Matched cue patterns will give one or more proposed CA for a given utterance segment. Native speakers can be asked to determine whether the correct label is among these, and if so, which label it is.
 - If the correct CA label is absent, its cue pattern set must be revised.
 - If the correct CA is present, natives should verify that substitution or paraphrase using alternative cue patterns for the correct CA would preserve communicative function in context; in case of failure, the cue pattern set should again be revised.
 - This cyclic revision is a training process, which repeats until little further revision is required.

Because CAs are defined in terms of surface cue patterns, it is possible to automate the first step of this procedure, the pattern matching step which proposes CAs. Such automation has the potential to greatly facilitate the training process.

Given the direct link between cue patterns and CAs, not only automatic CA analysis but automatic CA generation should be possible, and such generation can provide another important source of CA evaluation. Specifically, the suggestion is to use CAs as symbols in a representation which serves as a basis for generation.⁴ Since the purpose of CAs is precisely to symbolize what remains invariant in a range of cue patterns in context, paraphrases which are generated by varying the cue patterns during output should preserve the force or function in context. (In effect, such generation would automate the substitution step in the above evaluation procedure.) If the communicative function is not preserved, another cycle of revision is necessary to find distinctions or commonalities which were previously missed. Such an iterative discovery procedure is not new in the semantic area: it has been termed experimental semantics by [Hutchins 1971, 1975, Leech 1970, 1974, Mel'chuk and Zholkovski 1970] and others. The present proposal is to extend this sort of investigation to CAs as well.

We grant that the judgments concerning essential equivalence which form the core of this evaluation procedure are subjective. However, because they concern shared conventions regarding objectively observable cue patterns and contexts, a consensus (agreement) can be expected to emerge during repeated revision. This is the sense in which the proposed methodology is data-driven or corpus-based. Our present CA set represents such a consensus; but, as we have stressed, the cycle of revision will continue, and further refinement can certainly be expected.

Automatic selection of the correct CA is beyond the scope of this report. (See the Appendix regarding work in progress, however.) The current aim is to ensure that the correct CA label is among those suggested, and that the expressed communicative goal remains constant if alternative cue patterns for the correct CA are substituted in context.

2.1.7 What is not a CA?

We have been attempting to explain what a CA is. For contrast and comparison, let us discuss what it is not.

⁴IFTs are now used in exactly this way in ATR's ASURA speech translation system.

Some Communicative Goals are not Communicative Acts We have already stressed one negative several times: to guarantee a direct connection between CAs and surface language, our CA definition purposely excludes communicative functions which are not associated with conventional linguistic cues.

Thus our analysis is purposely limited to recognition of conventional surface discourse actions rather than deep intentions. We saw above that information about a bad dog, when provided without a conventional warning pattern, might or might not be interpreted as a deep warning, but in either case does not exemplify a WARNING CA. As a further example of the limitations of our approach, consider the fox in Aesop's story: when he says to the crow *Would you please sing a song for me?*, we can determine that he has made an ACTION-REQUEST, based on his use of the pattern "would you please *?", but we will not expect to discover his deeper plans to get the cheese which the crow is holding in his mouth. Similarly, *It's cold outside* would currently be analyzed as an INFORM, even though deeper plan-based analysis might show that the underlying or implied intent is to request the closing of a window.

This narrow focus is designed to enable experiments with computationally light discourse analyzers: initially at least, we can try to avoid the use of programs for deep plan recognition, in order to explore the limits of such processing.

Semantic Relations and Meta-relations are not CAs, Either Certain semantic relations have important roles in fostering global discourse coherence, since they can relate the propositions expressed by clauses, or sentences, or utterances. Relations concerning causation or temporal sequence are good examples. Since such relations usually can relate other semantic relations, they can be viewed as semantic meta-relations.

Because of this large-scale cohesive role, such semantic meta-relations are sometimes called "discourse relations" in the literature [see e.g. Hovy 1990]. We consider such relations to be important in analyzing discourse cohesion, but wish to distinguish them from Communicative Acts.⁵ Roughly speaking, semantic relations and meta-relations represent propositions, while CAs represent communicative goals concerning propositions: e.g. the intent to inform about them, or be informed about them, etc. And so, normally, a CA has a semantic relation, or proposition, as its argument.⁶ For example, (before a b) or (cause a b) would be treated as propositions, and could appear as arguments to CAs, as in (YN-QUESTION (before a b)) or (INFORM (cause a b)).

The problem is that in many CAs, some elements can be interpreted either as part of the communicative goals or as propositional content. For instance, we can consider establishing a DESIRE CA associated with the speakers' goal of informing about his or her own desired actions, with cue patterns like "I'd like/I want [VP, to infinitive]" or "[verb stem] たい" (*tai*). If we do so, we will represent the underlying form as something like (DESIRE (do speaker a)), where DESIRE is presumably a subcategory of INFORM, and the speaker's agent role in DESIRE is part of the CA definition. On the other hand, if we view desiring as a proposition, we will instead write something like (INFORM (desire speaker (do speaker a))), in which desire is a semantic relation like before or cause. Similar observations apply to BELIEVE and several other possible CAs which we considered.

Presently, we do not include DESIRE, BELIEVE, etc. as CAs. Since the relevant cue patterns and their interpretation in context are quite regular, there seems to be little reason to do so; but this choice is debatable, and we may change our mind in response to further evidence and experience.⁷ (By contrast, we treat cue patterns like "I'd like/I want you to *" as valid conventional patterns for

⁵However, discovery procedures similar to the ones proposed here for CAs are equally applicable to semantic meta-relations: see [Seligman 1991, 1994a].

⁶The opposite representation is also possible: a proposition can have a CA as its argument. The important point is that they are distinct but interrelated.

⁷There is a further reason for omitting DESIRE, BELIEVE, etc. from our current list: even if they were treated as CAs, they would presumably be subcategories of INFORM (as mentioned above), and thus perhaps too specific

ACTION-REQUESTS because (1) the usability of these cue patterns to make requests is not quite predictable for an “innocent” language learner, and (2) we have independent reason to establish this CA in association with many specialized cue patterns, including “please *” and “[verb-stem] て下さい” (*te kudasai*).

Sequence Controllers may be CAs, but ... The status of expressions which control conversational sequencing (lexical temporizers like “let me see”, transition cues like “now then”, even filled pauses, etc.) is somewhat debatable. Confusingly, these controllers are sometimes called “discourse markers” in the literature [Schiffrin 1987]. It may be possible to treat them as CAs. However, the communicative goals which they express seem to differ from the goals which most concern us here since they concern the *manner* of communicating rather than the *content*. In view of these differences, we postpone consideration of sequence controllers for later work.

However, note that some cue patterns resembling filled pauses receive the label EXPRESSIVE because they are judged to have an exclamatory element. In particular, Japanese “あ!” (*a!*) and “え!” (*e!*) are treated in this way.

Another special problem involves English “Okay” and Japanese “はい” (*hai*). These cues can express ACKNOWLEDGE and other relatively canonical CAs. But we judge that they can also function as sequence controllers, signaling that the speaker has completed part of a discourse and is ready to begin a new part. (Distinctive prosodic patterns may help to distinguish this use in the future.) For our current presentation, we supply CA labels in the first case, but not in the second.

Finally, Japanese “(です) ね” (*desune*) has a similar double role, since it sometimes marks the CAs CONFIRMATION-QUESTION, DO-YOU-UNDERSTAND-QUESTION, etc., but in other cases functions as a sequence controller with a temporizing or filler function similar to that of English “you know”. Again, no separate CA label is supplied in the second case.

Fragments are Sometimes CAs, but ... The proper treatment of fragments, too, is uncertain. On one hand, it seems clear that various sorts of elliptical responses serve as abbreviated INFORM CAs. In such cases, the hidden INFORM becomes visible if the missing words are added: *Who can do it? You can (do it).* or *Who is the President? Bill Clinton (is the President).*

On the other hand, other incomplete utterances arise from false starts or disfluencies: *Bill Clinton ... uh ... Several politicians have been involved in scandals lately.* Such fragments should probably have no independent CA labels.

Since contextual analysis will often be required to distinguish among the various fragment types, we presently label most fragments INFORM, for consistency. However, this is a temporary measure pending future research.

2.1.8 Ambiguity and Stages of Processing

As mentioned, a given cue pattern may belong to several cue pattern sets because it can indicate more than one CA. For example, based on surface clues only, and without considering the context, Japanese *hai* could (in our current analysis) receive any one of these labels: GREET, ACKNOWLEDGE, YES.⁸

Because the mapping between a surface-cue and CA is in general one-to-many, we intend to separate assignment of CAs into two stages: in Stage 1, we segment the corpus and assign all

for our present consideration.

⁸Of course, some ambiguity can remain even after contextual analysis. Even a native Japanese hearer, for instance, may sometimes be left unsure whether *hai* indicates a YES or just an ACKNOWLEDGE.

possible CA labels to each segment; in Stage 2, we select the correct label using information from surface context.

Our hope is that both stages can be performed automatically. (In fact, working programs for both stages now exist. See Appendix concerning Work in Progress.) In this document, however, we concentrate upon the preliminary label assignment of Stage 1. That is, we focus upon description of our labels and the mappings to surface cues which define them. Disambiguation techniques will be treated elsewhere (see again the Appendix).

In some cases of apparent ambiguity, several apparently competing descriptions can simultaneously be right. Consider the input *Could you please tell me your name*: on our current analysis, this utterance could be identified as ACTION-REQUEST based on the pattern "could you please *", but it could also be a WH-QUESTION based on the pattern "could you please tell me *". In general, since an utterance can fulfill several communicative goals at once, we assume that multiple CAs may sometimes be valid.

Further, as mentioned, it may be desirable to propose a hierarchy of CA specificity in which a WH-QUESTION is a subtype of ACTION-REQUEST – specifically, a request that the hearer provide a certain type of information. Automatic CA identification could then aim for a maximally specific identification to facilitate analysis of CA sequences. In the present example, expected responses to a WH-QUESTION would be more tightly constrained than expected responses for a more general ACTION-REQUEST.

Detailed consideration of such hierarchical analysis will appear in [Seligman 1994b]. For now, however, note that we have purposely omitted from our current set several CAs which we expect to analyze as relatively fine-grained subcategories of some present CA. For instance, we considered the possible CA EXPLAIN (with cue patterns like "* you see", "cause *", "it's just that *", and "[clause] わけ/のです" (*wake/no desu*)). However, we came to view this hypothesized CA as a subcategory of INFORM which is too fine-grained for our present level of experimentation. We may wish to include such subtypes in future CA sets.

2.1.9 The Need for More Abstract Levels of Analysis

Some plausible communicative functions involve relations among utterances. For instance, a conversational prompt, like a YN-QUESTION or an ACTION-REQUEST, usually leads to a corresponding response: one can reply to a question by giving the information which was requested; and, more generally, one can reply to a request by doing what was asked, promising to do so, etc.

The temptation to define "responsive" communicative acts like YN-QUESTION-RESPONSE or ACTION-REQUEST-RESPONSE is therefore quite strong.

However, we observe an important distinction: some potential "responsive" CAs have distinctive cue sets, and thus provide local surface cues to the responsive function in the utterance itself; while some have no such distinctive sets, so that the responsive role can only be recognized by contextual analysis.

With this difference in mind, we do allow the CA labels YES and NO as responses to YN-QUESTIONS, since we do observe distinctive surface pattern sets including cues like "that's right", and "そうです" (*sou desu*). By contrast, we do not allow any such CA label as ACTION-REQUEST-RESPONSE, because in our data we found no distinctive surface cue set associated with this responsive communicative function.

This distinction is rather subtle. If it is not maintained, however, circularities in CA definitions become a real danger: one tries to define the CA in terms of its relational role, but the relational

role must be described in terms of the local CAs.⁹ Let us state the distinction as a corollary of our basic CA definition:

- CA Definition, Corollary 1.

Labels which depend entirely on reference to the context are not permitted in our CA set. (On the other hand, “responsive” labels which do have distinctive cue pattern sets are indeed permitted.)

A further example may be useful. In our corpus, speakers very often repeated phrases, especially noun phrases, in order to confirm the information they had received.

Agent: The fare is 500 yen.

Client: 500 yen.

Agent: 500 yen.

One of our early analyses included a REPEAT-TO-CONFIRM CA, and used it to label the second and third utterances above. But we eventually abandoned this analysis, in view of the requirement that CAs be entirely specified by the structural properties of the utterance itself. The proposed responsive CA did not qualify, because it could only be recognized by noting that the utterance repeats a previous utterance. (At present we can perceive no distinctive prosody.) Thus we presently treat such utterances like other fragments: we label them INFORM with respect to their local cues, as a temporary measure pending future research. (See the discussion of fragments above.) However, we expect that the confirming function will be recognized at a more abstract level of analysis.

2.1.10 Future Discourse Analysis

Communicative functions which depend entirely on reference to the context do have a place in our overall design, but it is not at the level of CAs. Rather, we expect to employ such purely relational labels during Stage 2 at an abstract level of analysis where dialogue MOVES and EXCHANGES will be represented. (Compare [Stenströme 1994].)

This design sometimes leads to the establishment of rather vague labels at the utterance-local CA level. The current ACKNOWLEDGE label, in particular, (associated with cues like “はい” (*hai*) and “okay”) serves as a catch-all or place-holder for later contextual analysis: eventually, it is expected to realize various sorts of MOVES.

2.2 Multi-lingual CAs

A second major emphasis in our approach is the aim to examine the relationships between CAs in English and Japanese.

While compiling and revising our label sets, we have attempted to apply them to spontaneous dialogues in both languages. (Our corpora are described in [Loken-Kim et al 1993a, 1993b], and tagged sample dialogues in both languages are included Section 4 below.)

We include labels in our CA set if they are appropriate for *either* English or Japanese; and we find that in practice there is quite a large overlap between these very different languages, since the large majority of labels are actually suitable for *both* languages.¹⁰

⁹We are grateful to Toshiaki Iwadera for bringing these problems to our attention.

¹⁰This strategy for multi-lingual symbol definition has been described as a *polylingual*, as opposed to *interlingual*, approach in [Kay et al, 1994]. It was also suggested as a possible strategy for the transfer engine of the ASURA MT system in [Seligman et al 1993, Seligman 1993].

How do we know that a given CA has the *same* communicative function in the respective languages? Our current assumption is that we can confirm that CAs in different languages are at least closely similar by comparing their cue patterns and glosses: both should be plausible translations. Since the cue patterns are primary in defining CAs, their correspondence is the primary measure of a close CA correspondence across languages: a simple case is the use of “please *” and “[verb-stem] て下さい” (*te kudasai*), among other expressions, as respective cue patterns for ACTION-REQUEST.

We have given a primary place in our methodology to native judgments that several cue patterns perform essentially the same communicative function in context. While it may be questioned whether natives can actually make such judgments effectively, it is generally accepted that good translators can provide translations which preserve most essential communicative functions. Thus we might at some point consider using translation judgments as primary clues in discovering CAs. In this case, rather than use monolingual CA cue pattern sets to define monolingual CAs and then compare the CAs cross-linguistically as we now do, we might start by surveying *cue pattern translation sets*. We would then expect to discover only the CAs which were valid in all of the relevant languages. CA discovery would then be treated as a specialty within the area of example-based translation. This possibility remains as a topic for future research.

Only two CAs in the present set seem appropriate for only English or only Japanese. We now describe these two.

- **English but perhaps not Japanese: INSTRUCT**

Our corpus indicates that in English, when giving transportation directions, the patterns like “you [VP, present tense]”, “you will *”, and “you want to *” form a distinctive set for this discourse function. Thus we establish a CA label INSTRUCT. In Japanese, however, the patterns used for such instructions appear to compose a subset of the patterns used for making ACTION-REQUESTs: “[verb-stem] ていただく” (*te itadaku*), “[verb-stem] ていただいで” (*te itadaide*) etc., comparable to English “(would you) please * (for me)”. And so INSTRUCT, if it is a distinct CA in Japanese, would probably be analyzed as a relatively fine-grained subtype of ACTION-REQUEST. We now ignore several apparent fine-grained CA subcategories of this sort (like EXPLAIN, see above). Thus we will avoid the use of INSTRUCT in labeling Japanese dialogues for now, and leave future use as an open question. However, we do include this CA in our polylingual listing, with a warning.

- **Japanese but not English: CONFIRMATION-QUESTION-TO-SELF**

In both languages, a dialogue participant can use a CONFIRMATION-QUESTION to ask a partner to confirm a proposition: *This is the Kyoto-bound train, isn't it?*, この電車は京都行きですね? (*Kono densha wa, kyouto yuki desu ne?*) And in both languages, s/he can “think out loud” by addressing a CONFIRMATION-QUESTION to him/herself: [musing] *Hmm, I guess this would be the Kyoto-bound train, eh?*, うーん、この電車は京都行きだろうな。 (*Uun, kono densha wa kyouto yuki darou na*). In the first case, a confirmation or disconfirmation is expected; but in the second case the partner need not respond unless an error is likely. In Japanese, it appears that surface cues can distinguish these two types of CONFIRMATION-QUESTION in our corpus: we saw two examples of a confirmation question addressed by the speaker to himself, and in both, the informality of the questioning pattern (“[NP/adv/adj/verb] だろうな” (*darou na*)) marks the utterance as inappropriate for addressing the partner (for this purpose, the appropriate pattern would be “[NP/adv/adj/verb] でしょうね” (*deshou ne*)); and so the only remaining candidate for addressee is the speaker. In English, however, no such indications of the addressee are available in the text, so currently we have no local, cue-based ways of distinguishing confirmation questions to oneself from those addressed to others. (Future consideration of prosodic cues might add the necessary cues.) Thus we could not justify separate labels in a monolingual English treatment, but do make the distinction in our polylingual CA set.

Having stated our goals and principles, we now present our current list of Communicative Acts, with examples, explanations, and representative sets of surface patterns.

3 Tentative Lists of CA Labels: English and Japanese

For each tentative CA label in English or Japanese, we provide the following information:

- **Explanation or 説明.** A gloss giving the approximate sense of the label.
- **Illocutionary verbs or 遂行動詞.** Verbs or phrases which can be used to talk about the Communicative Act. See the above discussion of Leech's illocutionary verbs and their use as heuristics for CA discovery.
- **Pattern or パターン.** Cue patterns for the CA. May contain variable locations marked with * or constraint descriptions in [brackets].
- **Parameters or 与件.** Dimensions along which this CA can vary while still preserving its essential function. For example, an ACTION-REQUEST can vary in politeness.
- **Examples or 表層形.** Taken from the indicated EMMI dialogues in most cases. Some invented examples are also included to supplement attested examples. (Examples are omitted when cue patterns themselves provide sufficient illustration.)

3.1 English CA Labels

1. INFORM

explanation	S gives H information. INFORM can be elided, as in short answers to questions ("Who can do it?" "You can." "Who is the President?" "Bill Clinton.")
illocutionary verbs	assert, state, tell, let someone know, inform, indicate, go on record
pattern	Declarative clause syntax, with possible ellipsis and wide variety of aspectual and mood combinations. Fragments are presently considered INFORM; see discussion.
parameters	
examples	Conference office; You can travel a number of different ways; You're located right here (E-MM-01)

2. OFFER

explanation	S informs H that S is willing to perform some act A for person P (often H) if P wishes A. Often presented as a question about whether P wishes the act. The OFFER category often overlaps with PERMISSION-REQUEST for service roles ("May I help you?").
illocutionary verbs	offer; extend/make an offer
pattern	let me/us [VP]; why don't I/we [VP]; can/may I/we [VP] (especially Can/May I help you?); I/we can/could [VP]; how/what can/may/could I/we [VP] (especially How can I help you?, What can I do for you?). Many fixed expressions.
parameters	politeness; formality
examples	How can I help you? (E-TL-03); Let me pull up my map here (E-MM-02)

3. OFFER-FOLLOW-UP

explanation	As for OFFER, but used when some act has already been performed for person P to indicate that S is willing to perform another act. As for OFFER, this category often overlaps with PERMISSION-REQUEST. Usually fixed expression, as in examples below.
illocutionary verbs	
pattern	*anything else*?; *any other*? Clause syntax must be interrogative yn-question.
parameters	politeness; formality
examples	Is there anything else I can help you with? (E-MM-07); Can I help you with any other things? (E-MM-05)

4. INSTRUCT

explanation	S informs H how to reach H's goals (e.g. how to get somewhere) with authority or certainty, often by indicating the steps.
illocutionary verbs	give (you) instructions; tell/show (you) how
pattern	you [VP, present or future tense] (especially you (will) want/need to [VP], you (will) [VP], you are going to [VP]); imperative clause syntax.
parameters	politeness
examples	You want to go to the area in the middle here; At this station you'll want to change subways (E-MM-02)

5. SUGGEST

explanation	S informs H how H's goals can possibly be reached, but without authority or certainty. S typically says that some action is possible, or might be good, or asks rhetorically whether some action would be possible or good, etc.
illocutionary verbs	suggest, submit, propose
pattern	why don't you [VP]; you could/can/might [VP]; could(n't) you [VP]?; how about [VP, ing form]?
parameters	politeness, formality
examples	You can have them take you directly to the International Hall; You can catch a cab right there (E-MM-01)

6. ACTION-REQUEST

explanation	S informs H that S wants H to perform some action, but without authority.
illocutionary verbs	request
pattern	will/would/can/could you [VP]?; please [VP]; I/we would like you to [VP]; can/may I/we ask you to [VP]?
parameters	politeness, formality
examples	Can you give me a printout of this map (E-MM-02); Could you tell me please how to get from Kyoto Station to your conference center (E-TL-10)

7. YN-QUESTION

explanation	S informs H that S wants H to provide information about the truth or falsehood of a proposition.
illocutionary verbs	ask, request information
pattern	Yes-no interrogative clause syntax, with wide variety of aspect and mood; can/could/would you (please) tell me if/whether *; I'd like you to tell me if/whether *; I'd like to ask (you) if/whether *; may I ask (you) if/whether *; prosodic cues can also identify.
parameters	
examples	Do you see the map?; If I say this in English, will the taxi driver understand? (E-MM-01)

8. CONFIRMATION-QUESTION

explanation	S informs H that S wants H to provide information about the correctness or incorrectness of a quoted proposition.
illocutionary verbs	confirm
pattern	CLAUSE, (is that) right?; CLAUSE, tag question.
parameters	
examples	That's coins, right? (E-MM-06); The train goes east, doesn't it (fabricated example)

9. DO-YOU-UNDERSTAND-QUESTION

explanation	S informs H that S wants H to provide information about whether some information has been heard or understood, is satisfactory, etc. Can be a separate question ("Okay?") or a tag ("CLAUSE, okay?") following a plain INFORM, an INSTRUCT, etc.
illocutionary verbs	confirm one's understanding/satisfaction/approval
pattern	(CLAUSE) OK?; prosodic cues can also identify.
parameters	
examples	You just get on the Kyoto subway going north, OK (E-MM-07)

10. PERMISSION-REQUEST

explanation	S informs H that S wants H to provide information about whether H will give permission for someone (often S) to perform some action. Often used as ACTION-REQUEST or (YN/WH-QUESTION), as in the first example below.
illocutionary verbs	request permission/consent
pattern	May I/we [VP]?
parameters	politeness, formality
examples	May I ask your age?; May I speak with you? (fabricated examples)

11. WH-QUESTION

explanation	S informs H that S wants H to provide information about something other than truth or falsehood.
illocutionary verbs	ask, request information
pattern	wh-interrogative clause syntax with wide variety of aspect or mood; could/can/would you (please) tell me *; I'd like you to tell me *; I'd like to ask (you) *; may I ask (you) *?
parameters	
examples	What is the best way to get there?; How long a ride is it (E-MM-01)

12. YES

explanation	S informs H that the response to H's YN-QUESTION, CONFIRMATION-QUESTION, or DO-YOU-UNDERSTAND-QUESTION is affirmative.
illocutionary verbs	answer in the affirmative
pattern	yes; yep; yup; yea; uhum; (that's) right.
parameters	politeness
examples	

13. NO

explanation	S informs H that the response to H's YN-QUESTION, CONFIRMATION-QUESTION, or DO-YOU-UNDERSTAND-QUESTION is negative.
illocutionary verbs	answer in the negative
pattern	no; nope; naw
parameters	politeness
examples	

14. ACKNOWLEDGE

explanation	S informs H that information given by H has been heard, received, understood, etc. ACKNOWLEDGE is often used to maintain communication and be polite, without strong indication of understanding or agreement.
illocutionary verbs	acknowledge; recognize
pattern	I see; great; uhuh; (that's) fine; OK; you too; (all) right.
parameters	politeness
examples	

15. EXPRESSIVE

explanation	Fixed expression or interjection expressing emotion (e.g. being excited or impressed) or cognitive state. Signals of task-oriented errors like hitting the wrong key or dropping something are presently included in this category.
illocutionary verbs	exclaim
pattern	wow; whoops; oops
parameters	
examples	

16. GREET

explanation	Fixed expressions used for greeting.
illocutionary verbs	greet, welcome
pattern	hi, good morning, hello
parameters	politeness
examples	

17. FAREWELL

explanation	Fixed expressions used for terminating conversations.
illocutionary verbs	bid farewell, say goodbye, close
pattern	goodbye; bye
parameters	politeness
examples	

18. GOOD-WISHES

explanation	Fixed expressions used for expressing good wishes, usually before FAREWELL.
illocutionary verbs	extend/offer good wishes
pattern	have a good day/trip/etc.; take care
parameters	politeness
examples	

19. GOOD-WISHES-RESPONSE

explanation	Fixed expressions used for responding to good wishes, usually before FAREWELL.
illocutionary verbs	extend/offer good wishes in return
pattern	you too; same to you
parameters	politeness
examples	

20. THANK

explanation	Fixed expressions used for expressing thanks.
illocutionary verbs	offer/express thanks/gratitude/appreciation
pattern	thank you (very much); thanks (very much)
parameters	politeness
examples	

21. THANKS-RESPONSE

explanation	Fixed expressions used for responding to thanks.
illocutionary verbs	offer/express thanks/gratitude/appreciation in return
pattern	(you're) welcome.
parameters	politeness
examples	

22. APOLOGY

explanation	Fixed expressions used for apologizing. Often used to signal a speech error before a correction.
illocutionary verbs	offer/express apologies, apologize
pattern	(I'm) sorry.
parameters	politeness
examples	I'm sorry, (that's going east) (E-TL-02); (Oh wait) I'm sorry (that's east) (E-TL-03)

23. APOLOGY-RESPONSE

explanation	Fixed expressions used in responding to an apology.
illocutionary verbs	forgive; accept one's apology
pattern	that's all right
parameters	politeness
examples	

24. ALERT

explanation	Fixed expressions used by S to get H's attention. Compare VOCATIVE.
illocutionary verbs	get one's attention
pattern	excuse me; hey
parameters	politeness
examples	

25. CONFIRMATION-QUESTION-TO-SELF

explanation	Not a Communicative Act in English.
illocutionary verbs	
pattern	
parameters	
specific forms	
examples	

26. INVITE

explanation	S informs H that S wants H to inform S whether H wants to do something which is of benefit to both S and H, often together with S. Compare OFFER.
illocutionary verbs	invite, extend an invitation
pattern	would you like to [VP]?
parameters	politeness, formality
examples	Would you like to come over for dinner?; Would you like to see my etchings?

27. VOCATIVE

explanation	Name or role title used by S get the attention of a particular H. Compare ALERT.
illocutionary verbs	address one (as)
pattern	Proper name, or title such as nurse, waitress, sir
parameters	
examples	Fred! Waitress! (fabricated examples)

3.2 Japanese CA Labels

1. INFORM

説明	話し手は聞き手に情報をあたえる
遂行動詞	知らせる、言う
パターン	平叙文(断片を含む): 動詞(未然形)+う形, せ形, ; 動詞(連用)+て形 ¹¹ , ます形, ません形, 致す(いたす)形; 動詞(連体)+の(ん)形, わけ形; 動詞(終止)+と形; 名詞+です, に形; サ変動詞語幹, 動詞(連用)+でき形
与件	丁寧度, 動作主体, 情報の所在によって表層形が異なる
例	翻訳電話通信国際シンポジウムの会場へ行きたいんですけど

2. OFFER

説明	話し手が聞き手の利益になることを提案する
遂行動詞	提案する, 申し出る
パターン	平序文・疑問文での表現が可能: 動詞(連用)+ます形, たい形, お...致す形
与件	提案者の心的態度によって表層形が異なる
例	お名前をフロントで言って頂ければできるようにしておきますけれども

3. OFFER-FOLLOW-UP

説明	話し手が聞き手の利益になることを再度提案する
遂行動詞	再度提案する, 再度申し出る
パターン	平序文・疑問文での表現が可能: あと特に何かございませんか; あと他に何かございませんか
与件	特に/他になど副詞の変化あり
例	後は特に何かありますでしょうか

4. INSTRUCT

説明	Not a Communicative Act in Japanese?
遂行動詞	
パターン	
与件	
例	

5. SUGGEST

説明	話し手が間接的な方法で聞き手に意見を述べる
遂行動詞	ほのめかす
パターン	平序文・疑問文が可能: いかがですか, どうですか, と思うんですけども
与件	「ては」の文が前出する可能性あり
例	北側の出口を出て頂くと便利だと思うんですけども

6. ACTION-REQUEST

説明	話し手が聞き手に行為を要求する
遂行動詞	要求する, 求める
パターン	疑問文・命令文がおおい: 動詞(連用)+ていただく形, たい形, でき形, 願え形, ます形; 動詞(連体)+こと
与件	要求する態度によって表層形が異なる
例	そこをひだりにおれて頂きまして

7. YN-QUESTION

説明	話し手が聞き手に「はい」か「いいえ」で答えられる質問をする
遂行動詞	尋ねる, きく
パターン	疑問文: 動詞(連体)+んでしょうか, わけでしょうか, ことでしょうか, 名詞+ですか, 動詞(終止)+でしょうか, 動詞(連用)+ますか
与件	相手の意志を尋ねるか状況を尋ねるか, または話し手の心的態度で表層形が異なる
例	バス停は一つだけなんですか

8. CONFIRMATION-QUESTION

説明	話し手が聞き手に確認をする
遂行動詞	確認する, 確かめる
パターン	疑問文による: 動詞(終止形)+ということですか; そう+ですか; 動詞(連体)+んですか形, わけですな形; 名詞+ですな, ですよな形
与件	成行きとしての事実を確かめるか相手の意志を確かめるかによって表層形が異なる
例	会議の出席者の方ですよ

9. DO-YOU-UNDERSTAND-QUESTION

説明	話し手が聞き手に話し手の発話を理解したかどうか確認する
遂行動詞	尋ねる
パターン	動詞「わかる」の疑問文による: おわかりになりますか, わかりますでしょうか
与件	直接相手の理解を確かめるか間接的にかによって表層形が異なる
例	C: 分からないですけど蹴り上げるですね A: はいよろしいですか

10. PERMISSION-REQUEST

説明	話し手が聞き手に許可や了解を求める
遂行動詞	と許可をもとめる
パターン	疑問文・平叙文でへりくだった表現が多い: 動詞(連用)+せていただくことはできますでしょうか, お願いします
与件	直接相手の許可を求めるか間接的にかによって表層形が異なる
例	遅れて行きますがよろしいでしょうか

11. WH-QUESTION

説明	話し手が聞き手にいつ, どこ, だれが, いか, だれにを尋ねる
遂行動詞	尋ねる, きく
パターン	いつ, どこ, だれが, いか, だれにを含む疑問文
与件	
例	どのような字を書くんでしょうか

12. YES

説明	YN-QUESTION の肯定応答
遂行動詞	答える
パターン	はい, はいそうです
与件	決った表現
例	C: 国際シンポジウム会議事務局ですか / A: はいそうです

13. NO

説明	YN-QUESTION の否定応答
遂行動詞	答える
パターン	いいえ, いえ, いや
与件	決った表現
例	C: バスかなにかありますか A: いいえ歩いて行ける距離なんですけど

14. ACKNOWLEDGE

説明	聞き手が話し手に談話を継続させるためにあいづちをうつ
遂行動詞	とあいづちをうつ
パターン	はい, わかりました, そうですか, ええ, はあはあはあは, 承知しました
与件	決まった表現
例	C: あーめんどくさいですね / A: はいはあはあはあ

15. EXPRESSIVE

説明	感嘆詞をする
遂行動詞	と発する
パターン	え, しまった, えっ, あ
与件	比較的決まった表現
例	あ, すいません

16. GREET

説明	挨拶の言葉
遂行動詞	と挨拶する
パターン	もしもし, おはようございます, どうも
与件	決まった表現
例	A: あのこちらで予約させていただきますので C: あ, どうも

17. FAREWELL

説明	挨拶の言葉
遂行動詞	と挨拶する
パターン	失礼いたします, 失礼, 失礼します, 失礼致します, 分かりました, では, じゃ, それでは, よろしくお願ひします
与件	決まった表現
例	では, 失礼致します

18. GOOD-WISHES

説明	聞き手の幸福・安泰などを願う言葉
遂行動詞	とねぎらう
パターン	おきを付けて, おきをつけて, お気をつけて
与件	決まった表現
例	では, おきを付けて

19. GOOD-WISHES-RESPONSE

説明	幸福・安泰などを願う言葉に応える
遂行動詞	と応える
パターン	どうも
与件	決まった表現
例	どうも

20. THANK

説明	聞き手に感謝の言葉を述べる
遂行動詞	と謝す
パターン	ありがとうございました, 有難うございました, すいません, 有難うございます, どうもありがとうございます
与件	決まった表現
例	どうもありがとうございます

21. THANKS-RESPONSE

説明	感謝の言葉に応える
遂行動詞	と応える
パターン	どういたしまして, どうも
与件	決まった表現
例	どういたしまして

22. APOLOGY

説明	謝罪の言葉
遂行動詞	とあやまる
パターン	すいません, 失礼, *申し訳ないんですけども, お待たせ致しました, 失礼しました, お待たせしました, 長い時間取りまして, *申し訳ないんですけども
与件	決まった表現
例	お待たせしました

23. APOLOGY-RESPONSE

説明	謝罪の言葉に応える
遂行動詞	と応える
パターン	どういたしまして, はい
与件	決まった表現
例	A: 失礼致しました / C: はい

24. ALERT

説明	聞き手の注意を引く
遂行動詞	と注意を引く
パターン	すいません, 失礼, はい, もしもし, 失礼します, えっと
与件	決まった表現
例	ちょっと, すいません

25. CONFIRMATION-QUESTION-TO-SELF

説明	話し手が知り得たことを一人ごちる
遂行動詞	という
パターン	動詞(終止)+ なんだろうな
与件	
例	どこなんだろうなちょっとこれだけじゃわかんないんですけども

26. INVITE

説明	話し手が聞き手に何かを一緒にするよう誘う
遂行動詞	誘う
パターン	動詞(終止)+ ませんか
与件	
例	じゃこうしませんか

27. VOCATIVE

説明	名前などを呼びかける
遂行動詞	と呼びかける
パターン	固有名詞
与件	
例	あ, 野村様

4 Labeled Dialogues in English and Japanese

4.1 Two Dialogues in English

In the English transcriptions below, these conventions are used:

- Square brackets [...] denote filled pauses such as [um] and [ah].
- Curly brackets {...} denote speech which overlaps the speech of the other participant. +...+ can also be used to avoid confusing multiple uses of {...} when several instances of simultaneous speech occur close together.
- Round brackets (...) denote false starts.
- Slashes /.../ denote non-language sounds such as laughs or lipsmacks.
- "thi" and "e" indicate non-reduced pronunciations of *the* and *a*.

4.1.1 Conversation E3A

A : good morning (GREET) conference office (INFORM) {how} can I help you (OFFER)
BM : {/breath intake/} hi (GREET) I just arrived in Kyoto (INFORM) and I need directions to get to the International Conference Center (ACTION-REQUEST)
A : OK (ACKNOWLEDGE) (you're [ah]) you arrived in Kyoto (on the [s]) on the subway (INFORM)
BM : [y] [a?] no (NO) actually I just arrived on the Shinkansen (INFORM) I'm at Kyoto station (INFORM)
A : OK (ACKNOWLEDGE) [ah] you can get to thi International Conference Center by a number of ways (INFORM) either by bus taxi or the subway (INFORM) {how} would you like to travel (WH-QUESTION)
BM : {/breath intake/} [ah] I would prefer by the subway (INFORM)
A : OK (ACKNOWLEDGE) [ah] (to) to get to thi International Conference Center by the subway (you'll want to get on thi) let's see here you'll want to get on thi Higashiyama (INSTRUCT)
BM : [uh{um}] (ACKNOWLEDGE)
A : {line} and you can catch that at the second level floor platform (SUGGEST)
BM : OK (ACKNOWLEDGE) {second level} (INFORM)
A : {so (you're)} you're now on the first floor platform (INFORM) [ah] if you walk up [ah] some steps to the second level (INFORM)
BM : [uhu{m}] (ACKNOWLEDGE)
A : {th}ere's one platform (INFORM)
BM : [uhu{m}] (ACKNOWLEDGE)
A : {and} the subway leaves from there only north (INFORM)
BM : OK (ACKNOWLEDGE)
A : you'll wanna take the subway north to Sanjo station (INSTRUCT)
BM : OK (ACKNOWLEDGE)
A : at Sanjo station you'll need to change subways so get off at Sanjo [an] get on thi Keihan Kyotsu line (INSTRUCT)
BM : Keihan [kyotsu] (INFORM) how do you spell Kyotsu (WH-QUESTION)
A : [ah] Kyotsu is K Y O T S U (INFORM)
BM : OK (ACKNOWLEDGE)
A : so you'll take the Keihan Kyotsu line (INSTRUCT)
BM : {[mm]} (ACKNOWLEDGE)
A : {to} the second stop and that is Keage station (INFORM)

BM : [kee] (INFORM) [ho?] and how do you spell that (WH-QUESTION)
 A : that's K E A G E (INFORM)
 BM : {[m]} (ACKNOWLEDGE)
 A : {th}at's the second stop (INFORM)
 BM : OK (ACKNOWLEDGE)
 A : you'll get off at Keage station and walk west on Sanjo dori (INSTRUCT)
 BM : so west (INFORM)
 A : oh (ALERT)(I) wait (ALERT) I'm sorry (APOLOGY) that's east (INFORM)
 BM : OK (ACKNOWLEDGE)
 A : you'll be walking east /while laughing/ on Sanjo dori and you'll walk east to the very first street (INSTRUCT)
 BM : OK (ACKNOWLEDGE)
 A : you'll go left on the first street (INSTRUCT) it's a hairpin left (INFORM)
 BM : {OK} (ACKNOWLEDGE)
 A : it's a {very sharp} left (INFORM) and you'll walk about a half a mile (INSTRUCT)
 BM : [uhum] (ACKNOWLEDGE)
 A : the International Hotel is on the left in a great big large grey building (INFORM)
 BM : OK (ACKNOWLEDGE)
 A : OK (DO-YOU-UNDERSTAND-QUESTION)
 BM : OK (ACKNOWLEDGE) great (ACKNOWLEDGE)
 A : O{K} (ACKNOWLEDGE)
 BM : {tha}nk you {[mm]} (THANK)
 A : you're welcome (THANKS-RESPONSE) {have a go}od day (GOOD-WISHES)
 BM : you too (GOOD-WISHES-RESPONSE) {byebye} (FAREWELL)
 A : {bye} (FAREWELL)

4.1.2 Conversation E3B

A : Good morning (GREET) conference office (INFORM)
 BM : Hi (GREET) [um] I just arrived in Kyoto and I'm looking for the International Conference Center (INFORM)
 A : OK (ACKNOWLEDGE) are you at Kyoto Station? (YN-QUESTION)
 BM : Yes (YES)
 A : OK (ACKNOWLEDGE) [ah] we have three methods of travel to get to thi International Conference Center (INFORM) [ah] you can either go by subway bus or train (INFORM) what would you prefer (WH-QUESTION)
 BM : [um] I would prefer the quickest route (INFORM)
 A : OK (ACKNOWLEDGE) the quickest route would really be taking a taxi (INFORM)
 BM : {OK} (ACKNOWLEDGE)
 A : {and [uh]} I can call up the map here and [ah] we can take a look at how to get to the taxi stand for you (OFFER)
 BM : {OK} (ACKNOWLEDGE)
 A : {OK} so lets see (I'm gonna show you) this is a layout of the train station (INFORM)
 BM : {[um]} (ACKNOWLEDGE)
 A : {Where} you're located right now (INFORM)
 BM : OK (ACKNOWLEDGE)
 A : and you are on thi second floor platform right here (INFORM)
 BM : OK (ACKNOWLEDGE)
 A : and thi taxi stand is located to the north of the train station [ah] to thi [ah] east of the bus stop (INFORM) so to get there you take this passage way (INSTRUCT)
 BM : {[uh-huh]} (ACKNOWLEDGE)
 A : {across} the train station (INFORM) go through the door by the ticket office

across the street into the bus stop (INSTRUCT) and then there's a pathway that leads you right to the taxi stand (INFORM) that's {where} you can pick up a taxi (INFORM)

BM : {[oh]} OK (ACKNOWLEDGE) and is this all above ground? (YN-QUESTION)

A : [ah] yes (YES) the taxi stand is above ground (the platform) (INFORM) if you can see these steps here will take you down to thi first floor (INFORM)

BM : {[uh-huh]} (ACKNOWLEDGE)

A : {and then} down to the ticket office and up to the taxi stand (INFORM)

BM : OK (ACKNOWLEDGE)

A : {OK} (DO-YOU-UNDERSTAND-QUESTION)

BM : {great} (ACKNOWLEDGE) OK (ACKNOWLEDGE) and I can catch the taxi right in the front {there} (YN-QUESTION)

A : {Ye}s (YES) you can get the taxi right there (INFORM)

BM : OK (ACKNOWLEDGE) great (ACKNOWLEDGE)

A : Is there anything else I can help you with {cost} (OFFER-FOLLOW-UP)

BM : {[nnnn]} (ACKNOWLEDGE)

A : or anything (YN-QUESTION)

BM : [um] actually can you tell me (what) how much I can expect to pay (WH-QUESTION)

A : Yes (ACKNOWLEDGE) the taxi will run you about ten thousand yen (INFORM) and [ah] I'll show you (in relation to) [ah] [ah] in the Kyoto area where the International Conference Center is (INFORM)

BM : [mK] (ACKNOWLEDGE)

A : So you'll have an idea (where) where you're going here (INFORM) OK we're gonna look at another map now (INFORM)

BM : OK (ACKNOWLEDGE)

A : OK (ACKNOWLEDGE) ([th?]) thi International Conference Center is located right here (INFORM)

BM : [um-huh] (ACKNOWLEDGE)

A : and you'll be coming in on your taxi through this route then up here (INFORM)

OK (DO-YOU-UNDERSTAND-QUESTION)

BM : and that's ten thousand yen (YN-QUESTION)

A : {that's ten thousand} yen for +the+ ride (INFORM)

BM : {/breath/[?]} +[?]+ If I wanted to do something not quite as expensive could I take the subway (YN-QUESTION)

A : yes (YES) you can take the subway (INFORM) and I will call up another map (INFORM) and we can take a look at [ah] where you catch the subway at (INFORM) (the subway) lets see here OK we're going back to (thi) [ah] the map of the subway now (INFORM)

BM : [um-huh] (ACKNOWLEDGE)

A : [ah] on the second floor platform where you came in (INFORM)

BM : [um-huh] (ACKNOWLEDGE)

A : [ah] you can catch the same subway and (it goes) [ah] its thi Shinkansen Line (INFORM)

BM : [um-huh] (ACKNOWLEDGE)

A : and get on the subway (INSTRUCT) and it will take you to Sanjo Station (INFORM)

BM : O{K} (ACKNOWLEDGE)

A : {so you}'ll wanna take the subway north to Sanjo station (INSTRUCT) and (you can catch) you can catch the [ah] subway right there (SUGGEST)

BM : OK (ACKNOWLEDGE)

A : OK (ACKNOWLEDGE) now [ah] (when you're at [s] Sanjo Station) (I'll go look) [ah] lets see go back to the International Hotel (INSTRUCT) and I'll (give you) show you directions (INFORM) here you'll take the [ah] subway north to Sanjo Station (INSTRUCT) {here's San}jo (INFORM)

BM : {[uh-huh]} (ACKNOWLEDGE)

A : station (INFORM)

BM : 'K (ACKNOWLEDGE)

A : at that station you get off and change subways (INSTRUCT)

BM : OK (ACKNOWLEDGE)

A : and you'll get on thi Keihan-Kyotsu Line and go to the second stop on the Keihan-Kyotsu Line (INSTRUCT)
 BM : {OK} (ACKNOWLEDGE)
 A : {and that} will take you over here to ([kiage]) Keage Station (INFORM)
 BM : OK (ACKNOWLEDGE)
 A : At Keage Station you'll want to get off the subway and walk east on Sanjo Dori (INSTRUCT)
 BM : [uh-huh] (ACKNOWLEDGE)
 A : to the first street (INFORM) the first street take a left (INSTRUCT) walk about a half a mile (INSTRUCT) and the International Conference Center is on the left (INFORM)
 BM : OK (ACKNOWLEDGE)
 A : {OK} (DO-YOU-UNDERSTAND-QUESTION)
 BM : {/breath/} OK (ACKNOWLEDGE) great (ACKNOWLEDGE) {thank} you (THANK)
 A : {[o?]} you're welcome (THANKS-RESPONSE) {have a} good day (GOOD-WISHES)
 BM : {[mm]} you too (GOOD-WISHES-RESPONSE)
 A : good{bye} (FAREWELL)
 BM : {bye} bye (FAREWELL)

4.2 Two Dialogues in Japanese

4.2.1 Conversation J1A

A : はい (GREET) 国際会議事務局です (INFORM)
 MW : あ (EXPRESSIVE) すいません (APOLOGY) あの きょうの翻訳電電話通信国際シンポジウムの会場へ行きたいんですけど (INFORM)
 A : はい (ACKNOWLEDGE)
 MW : えっと どうゆうふうに行ったらいいんでしょうか (WH-QUESTION)
 A : えっと いまどちらにいらっしゃるんでしょうか (WH-QUESTION)
 MW : いま 京都駅降りたところなんですけど (INFORM)
 A : はい (ACKNOWLEDGE) えっと ですね
 MW : はい (ACKNOWLEDGE)
 A : えー 京都駅のあの烏丸町口を出て頂きます (ACTION-REQUEST)
 MW : 烏 (INFORM)
 A : バスに乗って頂くんですけども (ACTION-REQUEST)
 MW : 烏丸町ですか (YN-QUESTION)
 A : はいそうです (YES)
 MW : どういう字を書くんでしょうか (WH-QUESTION)
 A : あのー 烏に (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : 丸いって書くんですけど (INFORM)
 MW : はい (ACKNOWLEDGE) 烏丸町口 (INFORM)
 A : はい (ACKNOWLEDGE) ですね
 MW : 町口 (INFORM)
 A : はい (ACKNOWLEDGE) でそちらを出て頂いて (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : バスに乗って頂きます (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE)
 MW : バスはどちら行のバス乗ったらいいんですか (WH-QUESTION)
 A : えっと 京阪三条行です (INFORM)
 MW : 京阪三条行 (INFORM)
 A : はい (ACKNOWLEDGE)
 MW : はい (ACKNOWLEDGE)

A : えー 乗って頂きまして (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : あの終点の京阪三条まで行って頂きます (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : でそちよろしいですか (DO-YOU-UNDERSTAND-QUESTION)
 MW : はい (YES)
 A : はい (ACKNOWLEDGE) えー それからですね あのー 阪急電車 (INFORM)
 MW : 阪急電車 (INFORM) はい (ACKNOWLEDGE)
 A : あ (EXPRESSIVE) すいません (APOLOGY) まちがえました (INFORM)
 MW : えっ (EXPRESSIVE) はい (ACKNOWLEDGE)
 A : もしもし (ALERT)
 MW : あ (EXPRESSIVE) はい (ACKNOWLEDGE) はい (ACKNOWLEDGE)
 A : 京阪京津線にの乗り換えていただきまして (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE) 京阪京都線ですか (YN-QUESTION)
 A : はい (YES) 大津行の方に行って頂きます (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : で二駅であのー えっと これえー 蹴上の駅に到着するんですけども (INFORM)
 MW : え (EXPRESSIVE) すいません (APOLOGY) もう一回お願いします (ACTION-REQUEST)
 A : え (EXPRESSIVE) 蹴上です (INFORM)
 MW : 蹴上 (INFORM)
 A : はい (ACKNOWLEDGE)
 MW : どういう字ですか (WH-QUESTION)
 A : 蹴り上げるって書くんですけども (INFORM)
 MW : 蹴上 (INFORM)
 A : はい (ACKNOWLEDGE)
 MW : 蹴上駅 (INFORM)
 A : そうです (ACKNOWLEDGE)
 MW : はい (ACKNOWLEDGE)
 A : そちらの方を降りて頂きまして (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : えー 北側ですね 出口から出て頂きます (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE) 北側の出口 (INFORM) はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE) えー そちら北側の出口を出ていただいてから (ACTION-REQUEST)
 あのー もう少し大津の方に進行方向に (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : 歩いて頂きまして (ACTION-REQUEST)
 MW : えっと 大津って出口からどちら側になるんですか (WH-QUESTION)
 A : あの電車の進行方向です (INFORM)
 MW : 電車の進行方向 (INFORM) はい (ACKNOWLEDGE)
 A : はいそうです (ACKNOWLEDGE)
 MW : 電車上でんき電車沿いに道があるんですか (YN-QUESTION)
 A : あ (EXPRESSIVE) そうです (YES)
 MW : はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE) えー 約二メー二百メートルほど歩いて頂きまして (ACTION-REQUEST)
 Y字路がありますので (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : そちらを左側に曲がって頂きます (ACTION-REQUEST)
 MW : 二百メートルのY字路を (INFORM)
 A : はい (ACKNOWLEDGE)
 MW : えっと 右左 (WH-QUESTION)
 A : 左です (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE) で真っ直ぐ歩いて頂きましたら (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)

A : 約十分程で左手の方に会議場が見えてまいりますので (INFORM)
MW : はい (ACKNOWLEDGE) 左に曲がって約十分ですね (CONFIRMATION-QUESTION)
A : そうです (YES) はい (YES)
MW : はいわかりました (ACKNOWLEDGE) あ (EXPRESSIVE) どうもありがとうございました (THANK)
A : あ (EXPRESSIVE) どうも (THANKS-RESPONSE)
MW : はい (ACKNOWLEDGE)
A : おきをつけて (GOOD-WISHES)
MW : あ (EXPRESSIVE) はい (ACKNOWLEDGE) どうも (GOOD-WISHES-RESPONSE)
A : 失礼します (FAREWELL)
A : 失礼します (FAREWELL)

4.3 Conversation J1B

A : はい (GREET) 国際会議事務局です (INFORM)
MW : あ (EXPRESSIVE) ちょっと待って下さい (ACTION-REQUEST)
A : あ (EXPRESSIVE) はい (ACKNOWLEDGE)
MW : あ (EXPRESSIVE) もしもし (GREET)
A : もしもし (GREET) 国際 (INFORM)
MW : あ (EXPRESSIVE) すいません (APOLOGY)
A : はい (ACKNOWLEDGE)
MW : えっと きょうのあの翻訳電話通信国際シンポジウムの会議に
A : はい (ACKNOWLEDGE)
MW : 行きたいんですけど (INFORM)
A : はい (ACKNOWLEDGE)
MW : あの ちょっと場所がよくわかんないですよ (INFORM)
A : はい (ACKNOWLEDGE)
MW : それでちょっと場所を教えてくださいんですけど (ACTION-REQUEST)
A : はい (ACKNOWLEDGE) あのー 今現在どちらにいらっしゃいますか (WH-QUESTION)
MW : えっと 今京都駅降りたところなんです (INFORM)
A : 京都駅ですか (YN-QUESTION)
MW : はい (YES)
A : はい (ACKNOWLEDGE) 少々お待ち下さい (ACTION-REQUEST) あのー 今現在画面にでて
おります (INFORM) のが京都駅の地図なんですけれども (INFORM)
MW : はい (ACKNOWLEDGE)
A : あのーいまどちらにいらっしゃいますのでしょうか (WH-QUESTION)
MW : 今京都駅の裏側のここだと思っております (INFORM)
A : はい (ACKNOWLEDGE)
MW : はい (ACKNOWLEDGE)
A : えっと ここからでしたらあのー こちらの階段を上がって頂きまして (ACTION-REQUEST)
MW : はい (ACKNOWLEDGE)
A : ここずっと渡って頂きます (ACTION-REQUEST)
MW : はい (ACKNOWLEDGE)
A : で次の階段こちらが見えてきた時点で (INFORM)
MW : はい (ACKNOWLEDGE)
A : 左に曲がって頂きます (ACTION-REQUEST) 一番最初に出てきますこの交差点みたい
なところをですね (INFORM??)
MW : はい (ACKNOWLEDGE??)
A : 右に行ってください (ACTION-REQUEST)
MW : はい (ACKNOWLEDGE)
A : で右に行ってください (ACTION-REQUEST) でこの階段を降りて頂きます (ACTION-REQUEST) と
MW : はい (ACKNOWLEDGE)

A : あの烏丸町口に出てまいります (INFORM)
 MW : はあはあはあ (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE) でこぞっと出で頂きます (ACTION-REQUEST) たらこちらの方
 にあの一 バス停があります (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE) え (EXPRESSIVE) こちらのですねあの一 百二番系統のあのせん
 三条京阪行のですね
 MW : 三条京阪行ですか (YN-QUESTION)
 A : 市バスに乗っていただきます (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE) えーと バスのそうですねえー 終点まで乗って頂きます (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : と
 MW : 三条京阪行で終点まで (INFORM)
 A : はいそうですね (ACKNOWLEDGE) はい (ACKNOWLEDGE) ちょっと待って頂けますか (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : えっと 三条京阪からですね
 MW : はい (ACKNOWLEDGE)
 A : 京阪京津線に乗って頂きます (ACTION-REQUEST)
 MW : えっと 京阪京津線ですか (YN-QUESTION)
 A : はい (YES) えっと
 MW : 京津線 (INFORM)
 A : はい (ACKNOWLEDGE) 京阪京津線の大津行です (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : 乗って頂きます (ACTION-REQUEST) 二駅目で (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : あの蹴上という駅にまいります (INFORM)
 MW : 二駅 (INFORM)
 A : 二駅で (INFORM)
 MW : 蹴上 (INFORM)
 A : そうです (ACKNOWLEDGE)
 MW : はい (ACKNOWLEDGE) 蹴上ってゆうのどいう字を書くんですか (WH-QUESTION)
 A : えっと 蹴り上げるって書くんですけども (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : え (EXPRESSIVE) わかりますでしょうか (DO-YOU-UNDERSTAND-QUESTION)
 MW : ちょっと蹴る蹴り上げる (INFORM) ちょっとわかんないです (INFORM)
 A : えー ちょっとお待ち下さい (ACTION-REQUEST)
 MW : はい (ACKNOWLEDGE)
 A : このような名前の駅なんですけれども (INFORM)
 MW : あ (EXPRESSIVE) はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE)
 MW : はあはあはあ (ACKNOWLEDGE)
 A : え (EXPRESSIVE) そちらの方にえーと行って頂きますと (ACTION-REQUEST) ですね
 MW : はい (ACKNOWLEDGE)
 A : ちょっと待って頂けますか (ACTION-REQUEST)
 MW : あ (EXPRESSIVE) はい (ACKNOWLEDGE)
 A : え (EXPRESSIVE) いま 新しく出ました (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : あの地図のですね (INFORM)
 MW : はい (ACKNOWLEDGE)
 A : ここが蹴上の駅なんです (INFORM)
 MW : あ (EXPRESSIVE) はい (ACKNOWLEDGE)
 A : はい (ACKNOWLEDGE) でこちらの方が大津方面ですので (INFORM)
 MW : はい (ACKNOWLEDGE)

A : 進行方向に向かって歩いて頂きまして (ACTION-REQUEST)
MW : はい (ACKNOWLEDGE)
A : 最初のY字路を左に曲がって頂きます (ACTION-REQUEST)
MW : あはいはあはあ (ACKNOWLEDGE)
A : 約十分程直進して頂きます (ACTION-REQUEST) とあの左手の方ですねこちら (INFORM)
MW : はい (ACKNOWLEDGE)
A : こちらですね (CONFIRMATION-QUESTION)
MW : はあはあはあ (ACKNOWLEDGE)
A : こちらの方に会議場が見えてまいりますので (INFORM)
MW : はい (ACKNOWLEDGE)
A : はい (ACKNOWLEDGE)
MW : あ (EXPRESSIVE) わかりました (INFORM)
A : すぐわかりますので (INFORM)
MW : あ (EXPRESSIVE) どうも (THANK) 近くにもうほかにた高い建物とかはないんですか (YN-QUESTION)
A : や (EXPRESSIVE) ほとんど目印になるものはありませんし (INFORM) あの会議場の方がかなり大きな建物ですので (INFORM)
MW : はあはあはあ (ACKNOWLEDGE)
A : すぐにわかると思います (INFORM)
MW : はいわかりました (ACKNOWLEDGE)
A : はい (ACKNOWLEDGE)
MW : はい (ACKNOWLEDGE) どうも (THANK)
A : はい (ACKNOWLEDGE) おきをつけて (GOOD-WISHES)
MW : はい (ACKNOWLEDGE) 失礼します (FAREWELL)
A : 失礼いたします (FAREWELL)

REFERENCES

- Fais, Laurel and Gen-ichiro Kikui, 1991. *Determining Surface Form for Indirect Speech Acts in English*. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-0235.
- Iida, Hitoshi and Hidekazu Arita. 1992. "Natural Language Dialogue Understanding on a Four-layer Plan Recognition Model." In *Journal of Information Processing*, Vol. 15, No. 1, 1992.
- Hovy, E. 1990. "Parsimonious and Profligate Approaches to the Question of Discourse Structure Relations." In *Proceedings of the 5th International Workshop on Language Generation*, Pittsburgh, PA, June, 1990.
- Hutchins, W. 1971. *The Generation of Syntactic Structures from a Semantic Base*. London: North-Holland, 1971.
- Hutchins, W. 1975. "Subjects, Themes, and Case Grammar." *Lingua*, 35(2) (1975) 101-133.
- Ivir, V., and D. McMillan, T. Merz. 1973. "S-Relators." *Yugoslav Serbo-Croatian-English Contrastive Project*, 7 (1973), 22-64.
- Kay, M., J. M. Gawron, and P. Norvig. 1994. *Verbmobil: A Translation System for Face-to-face Dialog*. CSLI (Center for the Study of Language and Information) Lecture Notes No. 33.
- Knott, A. and R. Dale. 1992. *Using Linguistic Phenomena To Motivate A Set Of Rhetorical Relations*. Technical Report Rp-34, Human Communication Research Centre, University Of Edinburgh. Also in *Discourse Processes* 18(1) 1995, pages 35-62.
- Kume, Masako, and Gayle K. Sato, 1989. *A Descriptive Framework for Translating Speaker's Meaning: Towards a Dialogue Translation System between Japanese and English*. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-0118. Also in EACL-89.
- Leech, G. 1970. *Towards a Semantic Description of English*. Bloomington: Indiana University Press, 1970.
- Leech, G. 1974. *Semantics*. Baltimore, MD: Penguin Books, Inc., 1974.
- Leech, G. 1983. *Principles of Pragmatics*. London and New York: Longman.
- Loken-Kim, K. H., Yato, F., et al. 1993a. *EMMI-ATR Environment for Multi-Modal Interaction*. ATR Interpreting Telephony Research Laboratories, ATR-ITL Technical Report TT-IT-0081.
- Loken-Kim, Kyung-ho, Fumihiko Yato, Laurel Fais, Kazuhiko Kurihara, Ryo Furukawa, and Yoshihiro Kitagawa. 1993b. *Transcription of Spontaneous Speech Collected Using a Multimodal Simulator - EMMI, in a Direction-finding Task (Japanese-Japanese, English-English)*. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-IT-0029.
- Loken-Kim, Kyung-ho, Fumihiko Yato, Laurel Fais, Tsuyoshi Morimoto, 1994. "Linguistic and Paralinguistic Differences of Telephone-only and Multi-modal Dialogues." *Proceedings of ICSLP*, Yokohama, September 1994.
- Mel'chuk, I. and A. Zholkovski. 1970. "Toward a Functioning Meaning-Text Model of Language." *Linguistics*, 57 (1970), 10-47.
- Myers, J. 1990. *A Design for a Disambiguation-based Dialogue Understanding System*. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-189.
- Nagata, Masaki, Masami Suzuki, and Sachiyo Tsukawaki. 1993. 日英対話データベースへの発話行為タイプの付与方法に関する基礎的検討. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-0298.

- Sadanobu, Toshiyuki, Takayuki Yamaoka, and Hitoshi Iida. 1991. *Connections between Surface Sentence Forms and Speaker's Intentions in Cooperative Task-oriented Conversations; Classification of Communicative Acts Relied on Surface Forms*. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-0220.
- Schiffrin, Deborah, 1987. *Discourse Markers*. Cambridge: Cambridge University Press.
- Searle, J. 1969. *Speech Acts*. Cambridge: Cambridge University Press, 1969.
- Searle, J. 1975. "Indirect speech acts." In *Syntax and Semantics*, vol. 3, ed. P. Cole and J. Morgan. New York: Academic Press.
- Seligman, M. 1991. *Generating Discourses From Networks Using An Inheritance-Based Grammar*. Dissertation, Dept. of Linguistics, UC Berkeley.
- Seligman, M., M. Suzuki, and T. Morimoto. 1993. *Semantic-level Transfer in Japanese-German Speech Translation: Some Experiences*. Technical Report NLC93-13 of the Institute of Electronics, Information, and Communication Engineers (IEICE). May 21, 1993.
- Seligman, M. 1993. *A Japanese-German Transfer Component for ASURA*. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-0638.
- Seligman, M. 1994a. "Discovery and Format of Input Structures for Tactical Generation". In *Proceedings of the Seventh International Workshop on Natural Language Generation*. June 21-24, 1994, Kennebunkport, Maine.
- Seligman, M. 1994b. "Rhetorical Type Hierarchies and Disambiguation." Submitted to *Empirical Methods in Discourse Interpretation and Generation*, AAAI 1995 Spring Symposium Series, March 27 - 29, 1995, Stanford University, California.
- Stenström, A. 1994. *An Introduction to Spoken Interaction*. London and New York: Longman.
- Tomokiyo, Mutsuko and Tsuyoshi Morimoto. 1992. *Communicative Functions of Spoken Japanese and Its Meaning Interpretation on MT system*. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-0260.
- Tomokiyo, Mutsuko. 1993a. 語用論的分析に基づく自然発話の長文分割. ATR Interpreting Telephony Research Laboratories, ATR Technical Report TR-I-0019.
- Tomokiyo, Mutsuko. 1993b. *Transfert de la langue parlée japonais-anglais dans le système de traduction automatique ASURA*. TA-TAO UREF-AUPELF.
- Yato, Fumihiro, Kyung-ho Loken-Kim, Laurel Fais, and Tsuyoshi Morimoto. 1994. "Analysis of Utterances in Multi-modal Dialogues for Direction Finding Task." *The Transactions of the Institute of Electronics, Information and Communication Engineers*, Vol.J77-D-II(8), pp. 1475-1483.
- Wierzbicka, Anna. 1987. *English Speech Act Verbs*. Orlando, FL: Academic Press.

APPENDIX: WORK IN PROGRESS

EMMI Experiments

A series of experiments is being conducted in the ATR Environment for Multimodal Interaction (EMMI) in which speakers engage in a direction finding task and hotel reservation task in a variety of media: human interpretation, machine interpretation, by telephone, and via multimodal configurations. (Available media include on-screen video, a typescript interface, and on-screen graphics such as maps and documents which can be seen by both parties. Touch-screen technology permits drawing on maps which can also be seen by both participants.) The goal is to provide information on human-human and human-machine interaction which can inform the design of a multimodal component of a speech processing system. Part of the investigation focuses on the CAs used in various communication contexts. This research has required that all the dialogues from the first EMMI experiment, both English and Japanese, be labeled with CA labels. Based on this labeling, two papers have already reported on the use of CAs in the Japanese dialogues [Yato et al. 1994, Loken-Kim et al 1994], and an extended version of the second paper, intended for publication in a special collection of papers from the 1994 ICSLP conference, will include comparable statistics for the use of CAs in the English dialogues. As the analysis of the experimental work in EMMI continues, it is expected that CAs will be a useful tool for furthering our understanding of communicative interactions in various contexts.

Speech Synthesis Application

Alan Black of ATR-ITL Department 2 is investigating the various intonation contours used in different CA labeled phrases. He will extract rules that identify how accent types and boundary tones differ with respect to communicative act. The rules will be used to synthesize the appropriate prosodic tune for an utterance based on the CA of that utterance. The hope is that by giving the speech synthesizer the information that an utterance is, for example, an ACKNOWLEDGE, the machine will give that impression from its speech by producing an appropriate intonational pattern.

Discourse Context Analysis

Mutusko Tomokiyo of ATR-ITL Department 4 is analyzing Japanese dialogues for machine translation using CA labels. Spontaneously spoken Japanese consists of short clauses which are connected by auxiliary sequences, conjunctions, or adverbs. Utterances often contain repetitions, insertions, or revision fragments. To handle these linguistic phenomena, an utterance is segmented using pre-determined sequence cues into simple sentence-like units called *stars*. A CA label is then assigned to each star. Thus the final representation of an input utterance consists of a sequence of stars, each associated with a descriptive label representing its discourse function. Specialized functions are under development for selecting the appropriate CA label in context when pattern matching yields ambiguities.