On Representing Special Languages with FLBC: Message Markers and Reference Fixing in SeaSpeak

Steven O. Kimbrough[1] and Yinghui (Catherine) Yang[2]

[1] University of Pennsylvania, Philadelphia, PA 19104, USA,
kimbrough@wharton.upenn.edu
[2] University of California at Davis, Davis, CA, USA,
yiyang@ucdavis.edu

Abstract. SeaSpeak is "English for maritime communications." It is a restricted, specially-designed dialect of English used in merchant shipping and accepted as an international standard. This paper discusses, in the context of SeaSpeak, two key problems in the formalization of any such restricted, specially-designed language, viz., representing the illocutionary force structure of the messages, and formalization of such reference-fixing devices from ordinary language as pointing and use of demonstratives. The paper conducts the analysis in terms of Kimbrough's FLBC agent communication language.

1 Introduction

SeaSpeak is known as "English for maritime communications." It is the language of merchant shipping, a restricted, artificial, specially-developed, English-like language adopted in 1988 by the International Maritime Organization (IMO) of the United Nations for use in ship-to-ship and ship-to-shore communications. Part of the significance of SeaSpeak's success is that it demonstrates the value and use of specially-built artificial languages. The question then naturally arises of whether a designed special language might be fully formalized and used in machine-to-machine or human-to-machine communication. We have been intrigued by such possibilities and in consequence have been investigating SeaSpeak to this end.[1] In what follows we focus on two aspects of the larger programme of formalizing special languages:

- Illocutionary forces
- Reference fixing

These aspects of language, discussed in detail herein, are quite common. They occur in SeaSpeak, but also in nearly any special language that will be interesting. Our chosen vehicle of formalization, Kimbrough's FLBC, is also a

[1] [KLPY03,KY04]

special case. SeaSpeak and its ilk present an important test challenge for *any* agent communications language (ACL, of which FLBC is an instance). We shall present evidence in the form of analysis that indeed FLBC is adequate to the problems of representing illocutionary forces and fixing reference in SeaSpeak. The exercise and the lessons learned will apply in general to ACLs.

That is the overview. Details begin in the next section with some background on special languages.

2 Special Languages

Language enables communication. Languages inhibit it, for communication requires a common language and the cost of learning multiple languages raises an often unsurmounted barrier. Having a *lingua franca*, a general language known universally, would afford universal communication. At various times and places certain natural languages, such as Greek, Latin, Mandarin, and French, have approximated universal communication vehicles.

Today English in some form appears headed towards being the universal language of commerce and affairs. The fact remains, however, that universal proficiency in English is not around the corner. Further, even with universal fluency in English there are, and will always be, realms of discourse for which precise and accurate communication is required concerning specialized topics. It is not enough to have basic knowledge of English if the purpose of communication is air traffic control, navigation, law enforcement, and so on. In these and many other realms of discourse there exist specialized concepts and vocabulary that have to be mastered in the interests of efficient and effective communication. General fluency in English is not sufficient. Neither is it necessary.

Special languages can in principle be created that are relatively easy to learn and that are sufficiently expressive for particular purposes. They need be mastered only by a given community of interest. This idea has had an extensive history and considerable uptake, and it goes by a number of names. Including *planned languages*, the literature uses a number of other terms and recognizes a number of related concepts:[2]

> artificial languages, constructed languages (conlangs), invented languages, imaginary languages, fictional languages, etc., including universal languages, auxiliary languages, interlanguages or interlinguas, international languages; and also including logical languages, number languages, symbolic languages, etc. [Har02]

as well as others, including *restricted languages*, *designed languages*, and *sublanguages*.

[2] Many of these terms denote different, albeit related, concepts. We shall use *special language* as an umbrella term.

> Informally, we can define a sublanguage as the language used by a particular community of speakers, say, those concerned with a particular subject matter or those engaged in a specialized occupation.[Sag86, page 2]

Several artificial sublanguages have been fielded, and are successfully in use today.[3] Examples include AirSpeak, SeaSpeak, PoliceSpeak, and LinguaNet.[4] These languages were designed to be easily learned so that they can be spoken and heard effectively. Their employment and continued development today suggests they will be useful in the longer term. They employ a controlled, or restricted, vocabulary. It is typically true of sublanguages that their

> ... grammar contains additional rules not satisfied by the language as a whole. It also happens that some of the grammatical rules of the language as a whole disappear, i.e., do not apply, in a sublanguage. Since the sublanguage must satisfy the rules for the language, this disappearance is possible only if the rules are satisfied vacuously in the sublanguage, i.e., if certain word classes or well-formed sequences or transformations do not appear in the sublanguage. [Har68, page 154]

Telegraphic languages, yet more austere forms of sublanguage, are also widely in use and readily display the simplified, constrained grammar characteristic of sublanguages. We note that *telegraphic language* and *telegraphic speech* are also terms of art in the field of child development, and it is here that the terms obtained their original meaning.

> When children initially produce grammar, their language often sounds rather like the abbreviated language of telegrams ("Daddy gone," "Mummy shoe." "See big car"). This is why, in the past, this type of early output was referred to as telegraphic speech. At this stage, toddlers omit indefinite and definite articles, as well as prepositions and the like. They also leave out morphemes like plural "s," progressive "ing," and possessive "'s." [KK01, page 94]

Telegraphic languages are not unknown among adults. Fitzpatrick et al. [FBH86] present a particularly clear case study of a telegraphic language used in the U.S. Navy. The stylization apparent in examples from this language—e.g., "72 manhours expended," "Stock requisition shipped," "Work request submitted," "Improper repair work performed," "No parts required" [FBH86, page 45]—will be familiar to the reader.[5]

[3] We shall not discuss various more ambitious efforts to develop general-purpose universal languages. Esperanto is perhaps the most well-known candidate language. For relevant background see [Lar85,Mac30,Ogd38,Ric43,Swa80]; also, Harrison [Har02] has put together a very useful bibliography.

[4] See [Ben03], [Joh98], [Joh02], [Pro03], [Lin03], and [J$^+$93] for an overview.

[5] Portions of this section contain a revised version of material appearing in [KLPY03].

3 Two Problems

Can special languages—especially artificial languages, sublanguages, and telegraphic languages—be formalized so that machines may productively conduct inferences using them? The question is significant both theoretically and and for applications, as has been noticed, e.g.,

> An interesting suggestion which could have widespread applications is that particular subdisciplines do in practice use a limited set of grammatical structures as well as a restriced vocabulary: a sublanguage or metalanguage, easily comprehended by those within the subdiscipline but foreign to the layman. Because of the limited number and specialized nature of the grammatical structures found it becomes possible to apply content analysis techniques to texts within the subdiscipline with success consistently. The particular subdiscipline illustrated is pharmacology, but it seems likely that the approach would be valid in any of the 'hard' sciences with a clearly defined vocabulary (jargon!) and generalized methodology. [Fos82, pages 51–2]

(See also [Sag75].) Theoretically, the question presents an apt challenge for ACLs (agent communication languages), including the various projects to create FLBCs (formal languages for business communication), and the various XML representation efforts. Can the ACLs adequately represent a given artificial language? If not, how might they be improved? What does formalization of artificial languages tell us about requirements for ACLs? From a practical point of view, formalization could afford human-machine and machine-human communication, including language translation and error detection, as well as machine-machine communication, with its attendant possibilities of reducing time and labor costs. Perhaps of most immediate use, formalization and structuring present opportunities for automated recovery and discovery of information.

These are large and fascinating questions, which succinctly put the context for the results reported in this paper. We essay here to make a modest, yet discernible, contribution to the advance on them. We shall examine one (informal) artificial language—SeaSpeak—and one variety of ACL (agent communication language), Kimbrough's FLBC, based on event semantics, thematic roles, and disquotation of propositional content. (For background on FLBC see in this volume "A Note on Modeling Speech Acts as Signalling Conventions" [JK04] and "Practical Contract Storage, Checking, and Enforcement for Business Process Automation" [AEB04].[6]

Specifically, SeaSpeak is a notable example of an artificial sublanguage, which is established and used successfully, and which might benefit from formalization. We have been investigating the prospects for such a formalization

[6] Other references for FLBC include [Kim90], [Kim99], [KM97], [KT00], [Kim01], [Kim02], [KLPY03], and [KY04].

(see [KLPY03,KY04]), with positive results, both for SeaSpeak in particular and for artificial sublanguages more generally. In what follows, we address in detail two key technical issues whose resolution is essential for any programme of formalizing communication among artificial (and human) agents in other than very restricted domains. The two issues are:

- The speech act structure of SeaSpeak messages.
- Dynamic reference fixing in SeaSpeak messages.

Central to speech act theory[7] is the distinction between the *illocutionary force* and the *propositional content* of an utterance. This distinction, commonly thought to originate with Austin [Aus62], goes back at least to Charles Sanders Peirce, in the nineteenth century.

> Like other philosophers of thought and language, Peirce distinguished the force of an utterance from its propositional content. A proposition can be 'affirmed, denied, judged, doubted, inwardly inquired into, put as a question, wished...' *Assertion* of a proposition involves 'the deliberate exercise, in uttering a proposition, of a force tending to determine a belief in it in the mind of the interpreter' .... Assertion involves 'taking responsibility' for the truth of the proposition. [Hoo02, page 62]

In any event, under the perspective of speech act theory, which is widely accepted and which we accept,[8] every utterance (in any language) may be analyzed as having an $F(P)$ structure: an illocutionary force, F, is applied to a propositional content, P. In general, illocutionary forces and propositional contents are in a many-to-many relationship. One force—e.g., asserting, directing, promising—may be applied to many propositional contents—e.g., 'I will arrive tomorrow', 'The tide will come in at 6 p.m.', 'An act of nuclear terror will strike New York City within 10 years'. Similarly, one propositional content may be the object of several different illocutionary forces. One may assert it, deny it, promise it, and so on. If this most basic tenet of speech act theory is correct, then it should be possible to recognize speech acts in SeaSpeak and to say something in general about their structures. In fact, as we shall see, SeaSpeak presents a happy confirmation of the $F(P)$ thesis and, as we shall argue, the SeaSpeak illocutionary forces (called *message markers* in SeaSpeak) may be aptly represented in FLBC.

The problem of dynamic reference fixing arises outside the perspective of speech act theory. We present and discuss the problem in depth in §6. Briefly, it is this. In communicating we wish to talk about things, about

[7] Classically, [Aus62], [Lev83], [Sea69], and [SV85]; [BH79] is useful; [LAP04], and [KM97] present application-oriented summaries.

[8] The Language/Action Perspective community has been holding workshops and producing papers for some years now, promoting and articulating applications of speech act theory. See [LAP04].

shoes, ships, sealing wax, cabbages, kings, kinghood, numbers, beliefs, and even things that do not exist such as unicorns. To talk about any thing we need to make reference to it. If the thing has a proper name, this is relatively unproblematic. Usually, however, we make reference to something with a *the*-expression—as in *The cat is on the mat*—or some similar device. The problem of dynamic reference fixing for an ACL is the problem of formalizing such expressions as *The cat is on the mat* in such a way that reference is successful and recoverable by the addressee of the utterance. SeaSpeak is replete with such expressions, particularly *the*-expressions, and so provides an opportune context in which to tackle this problem.

All of this requires a bit of background on SeaSpeak. To that now.

4 Background on SeaSpeak

SeaSpeak was developed and deployed in consequence of vastly increased shipping during the 1960s and 1970s. At the same time, the distribution of nationalities of ships' officers gradually changed from roughly 80% English-speaking and 20% other to roughly 80% other and 20% English-speaking. The need for regularization of practices in one language and the training of officers in its use was therefore agreed, and English, already the language of civil aviation, was chosen by the IMO.

During 1982–1983, SeaSpeak was created by specialists in maritime communications and applied linguistics [Joh02]. SeaSpeak is a system for speech communication, and it is intended for use in situations where it is essential that communication should be as clear, brief and accurate as possible.

Like SeaSpeak, Airspeak and Policespeak are also special purpose systems for speech communication among targeted users. They are the special languages of command and control where the utterances you make affect something far away—applying to communication between ships, for air traffic control and in police operations. Edward Johnson, Senior Fellow of Wolfson College, University of Cambridge, U.K., has been a pioneer in the field of operational and communication languages. Johnson was responsible for formulating an international language for maritime communication, SeaSpeak (1982), an air traffic pilot training communication program, AirSpeak (1986), and a restricted operational language and set of procedures for police communication, PoliceSpeak (1987).

SeaSpeak regulates ways of speaking and ways of establishing a conversation. It defines a technical vocabulary. All messages begin with a *message marker* that indicates the nature of what follows, such as advice, information, instruction, intention, question, request, warning, or a response to one of these. The definitive reference for SeaSpeak is the *SeaSpeak Training Manual* [WGJS88], upon which we draw for our analysis and formalization.

From a formalization perspective the central concept in SeaSpeak is the message marker. The manual has this (and not much else) to say about message markers in general.

> Maritime messages transmitted over VHF should be short, accurate, and relevant. Furthermore, messages should be transmitted in language simple enough for a non-native speaker of English to comprehend without difficulty.
> One useful means of making the language simpler is to indicate, at the beginning of a message, what sort of message it is going to be. Thus, if a question is going to be asked, the speaker simply says the word 'QUESTION' before the question itself. Similarly, if a piece of advice is going to be given, the speaker says the word 'ADVICE' in advance of his message. There are just severn of these *Message Markers* and after a little practice, learners should experience no difficulty in using them.
> These *Message Markers* have another function: that of imposing order on the conversation, since each message marked in this way requires a reply correspondingly marked (even if that reply is nothing more than an acknowledgement of the message received). This procedure helps to ensure that:
> 1. messages do not become confused with each other
> 2. each message is dealt with in turn
> 3. a participant receiving a reply knows which message is being replied to.
>
> [WGJS88, page 96]

SeaSpeak has only seven markers (with a mirroring reply-marker in each case). The seven are [WGJS88, pages 96–7]:

1. Information (Information-Received)
2. Warning (Warning-Received)
3. Intention (Intention-Received)
4. Request (Request-Received)
5. Advice (Advice-Received)
6. Instruction (Instruction-Received)
7. Question (Answer)

SeaSpeak's message markers are, as we shall analyze them, essentially speech act operators or illocutionary force indicators. SeaSpeak sentences have, we shall argue, the $F(P)$ structure posited by speech act theory. The Fs of SeaSpeak are its message markers. The content that they govern is simple in form, although not entirely specified and closed. Here is a summary from the training manual.

1. SEASPEAK messages are formed entirely from words within the English language.
2. The total vocabulary used in SEASPEAK comprises 3 kinds of words and expressions:
   (a) **The vocabulary of 'general' English.** Knowledge of the non-specialized vocabulary of English is assumed, and so it is not listed in the SEASPEAK Vocabulary.
   (b) **Words in general maritime use.** These words occur frequently in maritime communications, and are listed in Section I, as Categorised General Maritime Vocabulary.
   (c) **Words in specialised maritime use.** These words and expressions may occur only rarely in general maritime use, but frequently in particular circumstances or for specific communication subjects. They are listed in Section II under the Major Communications Subjects.

[WGJS88, page 160]

Item (2a) presents a particular challenge to formalization efforts. Just what is the scope of 'general English'? Examples are useful. Here and in the sequel we will use *italic font* for messages in SeaSpeak. The following examples are from [WGJS88, page 97].

1. *QUESTION: What is your ETA at the dock entrance?*
2. *INSTRUCTION: Go to berth number: two-five.*
3. *ADVICE: Anchor, position: bearing: one-nine-four degrees true, from Keel Point distance: one mile.*
4. *REQUEST: Please send, quantity: five acetylene cylinders.*
5. *INFORMATION: The pilot is waiting now, position: near buoy number: two-six.*
6. *WARNING: Buoy number: two-five and buoy number two-six are unlit.*
7. *INTENTION: I intend to reduce speed, new speed: six knots.*

We cannot fully address here the use of 'general English' in SeaSpeak. Instead, we focus on formal analysis of the seven message markers and on dynamic reference fixing, particularly use of *the*-expressions, as seen in the examples above.[9]

5 Prototype Example: INFORMATION

Our purpose in this section is to present a prototype, or illustrative, example of representing a SeaSpeak sentence in FLBC. Consider then the simple SeaSpeak sentence:

[9] Portions of this section contain a revised version of material appearing in [KLPY03].

SeaSpeak Sentence 1 *INFORMATION: No vessels are at the anchorage.*

We'll begin by analyzing the content of the simple sentence, *No vessels are at the anchorage,* deferring briefly discussion of the illocutionary force indicator, *INFORMATION.* Thus,

SeaSpeak Sentence 2 *No vessels are at the anchorage.*

Note that *the anchorage* is a referring expression whose meaning cannot be recovered from the bare sentence. To begin, we assume for the sake of the example that a definite anchorage has been identified, called (having proper name) *XBar-Harbor-B.* Later we will discuss at length our analysis of and approach to such expressions as *the anchorage* and *the harbor.*

The first step in formalizing a given SeaSpeak sentence is to convert it to more transparent forms, while remaining in natural (informal) language. The alternative forms are called *stylistic variants.* They are intended to be adequately similar in meaning (for the purposes at hand) to the original sentence, while also being more transparent for purposes of formalization. In this example we eliminate the referring *the*-expression in favor of its corresponding proper name. The first stylistic variant is thus:

SeaSpeak Sentence 3 *No vessels are located at XBar-Harbor-B.*

Notice that sentence 3 is arguably a clearer version of sentence 2, since the answer to the question Which anchorage? may be read off by inspection. We need one further transformation. The result is rather stilted, but fits the perspective of event semantics as deployed by FLBC.

SeaSpeak Sentence 4 *(1) There is a state of being a vessel, s_1, at this time. (2) Nothing now in the vessel state, s_1, is located at XBar-Harbor-B.*

SeaSpeak Sentence 4 is a well-structured stylistic variant of a SeaSpeak *sentence*; it does not constitute a complete SeaSpeak *message*, if only because the speaker and addressee are not identified. Figure 1 contains a complete representation (with comments) of a SeaSpeak *message.* Note that it is divided into three parts:

1. Reference-fixing
2. Presupposition
3. Message body

Points arising:

1. Fully articulated the message body is

   SeaSpeak FLBC Sentence 2 $information(e_1) \wedge Speaker(e_1, s) \wedge Addressee(e_1, a) \wedge Cul(e_1, now) \wedge Object(e_1, \lceil \forall x(In(x, s_1) \rightarrow \neg Located(x, a_1)) \rceil)$

1. Reference-fixing
   s = the speaker
   a = the addressee
   a_1 = *XBar-Harbor-B* (the name of a particular anchorage)
2. Presupposition
   (a) There is a state, s_1, of being a vessel at this time:
   $vessel(s_1) \wedge\ Hold(s_1,\ now)$
3. Message body
   (a) Propositional content.
   $\forall x(In(x, s_1) \rightarrow \neg Located(x, a_1))$
   (b) Illocutionary force.
   Further, the utterance is in the class INFORMATION, which indicates its illocutionary force. SeaSpeak FLBC Sentence 1 represents the INFORMATION force for the utterance to hand, with ϕ serving as a place-holder for the propositional content.
   SeaSpeak FLBC Sentence 1 $information(e_1) \wedge\ Speaker(e_1, s) \wedge\ Addressee(e_1, a) \wedge\ Cul(e_1,\ now) \wedge\ Object(e_1, \lceil\phi\rceil)$

Fig. 1. Basic example for FLBC representation of SeaSpeak INFORMATION messages

2. Rendered back into still stilted English, the message body says that e_1 *is an INFORMATION utterance event, occurring now, whose speaker is s, whose addressee is a, and whose propositional content is that no vessel is located at* a_1.
3. SeaSpeak FLBC Sentence 2 is a formalized representation of SeaSpeak Sentence 1 and its stylistic variants. None of these sentences is by itself a complete message. Abstracting Figure 1 to a more general template, we will use the message body item (#3) for representing SeaSpeak sentences.
4. The reference-fixing item (#1) of Figure 1 identifies the speaker, the addressee, and any other proper nouns (here, *XBar-Harbor-B*) required for the message. Strictly speaking, this section could be eliminated and the logical names, e.g., s, could be replaced throughout with their associated proper names, e.g., *Land's End Radio*. It is convenient, however, to retain this section.
5. Use of the presupposition section will be discussed in the sequel. Here, it is used to state the presupposition that there is a vessel state, in which something might potentially be in. Again, this is not strictly speaking necessary. One might drop this presupposition in favor of an expanded (but not equivalent) SeaSpeak FLBC Sentence:

   SeaSpeak FLBC Sentence 3 $information(e_1) \wedge\ Speaker(e_1, s) \wedge\ Addressee(e_1, a) \wedge\ Cul(e_1,\ now) \wedge\ Object(e_1,\ \lceil\forall x \forall s_1(vessel(s_1) \wedge\ Hold(s_1,\ now)\ \wedge\ In(x, s_1) \rightarrow \neg Located(x, a_1))\rceil)$

   (Note that now s_1 is being used as a variable, not a constant.)

In sum, the pattern in evidence here represents a SeaSpeak *message* as a structure, consisting of a reference-fixing section, a presupposition section, and a message body section, the latter being used to hold representations of particular SeaSpeak *sentences*. We devote the balance of this paper to articulating this pattern with further examples and in response to certain problems. The bulk of the problems we address fall into the category of how reference (to external objects) may be fixed in an automated or semi-automated system. We finessed that issue for the sake of the example in the present section. We now address it directly.

6 Problems of Reference Fixing

Ordinary discourse, and certainly SeaSpeak discourse, is replete with references to entities of various sorts, including particular objects, places, times, sounds, happenings, lengths, performances, fictional characters, geographical objects, and social entities (cf., [Jac02, pages 300–3]). Consider a simple example from ordinary language:

- *The cat is on the mat*

Both *the cat* and *the mat* are referring expressions. Which cat? Which mat? These references have to be adequately fixed if the hearer is to understand the sentence. In human–human discourse we have a number of devices we employ, often automatically and without deliberation or even awareness, so that both speaker and hearer will know what the sentence is about. Thus, e.g., *the cat* might harken back to an earlier part of the conversation, or a cat might be pointed to (in some way or another) by the speaker at the time of sentence utterance in such a manner that the speaker has good reason to believe that the hearer has "gotten the point." Similarly, *the mat* might refer by conventions of discourse to a particular mat identified earlier in the conversation.

The subject of how reference is fixed in ordinary discourse is a large and fascinating one. Our concern here is the more limited one of how this might be achieved in machine–to–machine (or process–to–process) discourse using a formal language. How might one process point to something in the world and have the other process be able to discern the object of the reference?[10] We turn now to *the*-expressions, which shall constitute the focus of our attention on these matters.

The word *the* is not always unproblematic, especially in the present context. Consider these typical SeaSpeak sentences:

[10] Note we say *discern* rather than *understand*. Computer processes supporting communication need not 'understand' messages, whatever that murky term may mean. If, however, a message is sent referring to an object, the receiving process will need to pick it out, distinguish it, discern it.

SeaSpeak Sentence 5 *QUESTION; What is your ETA at the dock entrance? [WGJS88, page 97]*

SeaSpeak Sentence 6 *INFORMATION: No vessels are at the anchorage. [WGJS88, page 101]*

SeaSpeak Sentence 7 *INFORMATION: The icebreaker intends to assemble the convoy at time: zero-five-three-zero GMT. [WGJS88, page 101]*

SeaSpeak Sentence 8 *INFORMATION: The casualty is approximately, position: North distance: three miles from you. [WGJS88, page 101]*

The first uses *the dock entrance* to refer to a particular dock entrance. Which one? The second sentence uses *the anchorage* in referring to a specific anchorage. Again, which one? In order for communication to be entirely successful the addressee must be able to ascertain, or pick out, the particular dock entrance or anchorage referenced by the speaker. How does this happen? How could this be brought about in an automated (or partially automated) context, as is under discussion here?

If the referents—the dock entrance and the anchorage—are given proper names—say Dock Entrance A and Anchorage D—and these are known and agreed upon prior to communication, then the problem is greatly simplified. Instead of speaking as above, SeaSpeakers could talk as follows.

SeaSpeak Sentence 9 *QUESTION: What is your ETA at Dock Entrance A?*

SeaSpeak Sentence 10 *INFORMATION: No vessels are at Anchorage D.*

This is exactly what we did in §5, naming *the anchorage* with *XBar-Harbor-B*. More generally, it is certainly possible to construct antecedently a table of proper names and their meanings, and to include this table in the working definition of a communications language, whether it be informal, as in SeaSpeak, or formal as in an FLBC. In the case of Maritime communications, listing all ships, harbors, lighthouses, and officers in this way would no doubt be useful. Practically, however, the approach has its limitations. Harbors and lighthouses may be more or less permanent, but ships and officers come and go. Ships may be renamed. Officers may fall ill during a voyage and new officers be declared. How will the list be maintained, how will the language be updated, and how will the updates be promulgated? Making matters worse, many of the things that need to be discussed are ephemeral: storms and other weather patterns, accidents and other incidents at sea, and so on. It is not a practicable possibility to name antecedently such referents as *the storm* in

SeaSpeak Sentence 11 *QUESTION: When will the storm arrive?*

In natural language—spoken and written, including sublanguages such as SeaSpeak—the complications of reference fixing are manifold. This creates enormous difficulties for information retrieval, as is well known and as is so charmingly described in the following passage from a famous empirical study.

> Sometimes we followed a trail of linguistic creativity through the database. In searching for documents discussing "trap correction" (one of the key phrases), we discovered that relevant, unretrieved documents had discussed the same issue but referred to it as the "wire warp." Continuing our search, we found that in still other documents trap correction was referred to in a third and novel way: the "shunt correction system." Finally, we discovered the inventor of this system was a man named "Coxwell" which directed us to some documents he had authored, only he referred to the system as the "Roman circle method." Using the Roman circle method in a query directed us to still more relevant but unretrieved documents, but this was not the end either. Further searching revealed that the system had been tested in another city, and all documents germane to those tests referred to the sysem as the "air truck." At this point the search ended, having consumed over an entire 40-hour week of on-line searching, but there is no reason to believe that we had reached the end of the trail; we simply ran out of time. [BM85]

Our problem is akin to, yet distinct from, the information retrieval problem. The latter aims at finding multiple references to a common object. The question before us is how a speaker may establish reference to an object in such a way that the addressee will be able to distinguish, pick out, establish her own reference to the object.[11]

In seeking to formalize and automate it will be helpful to examine how reference is fixed in natural language, for the devices employed there, or something like them, may prove useful in our context. Linguists and philosophers have discerned a number of reference-fixing devices in natural language, in addition to the use of proper names, already noted. For present purposes, these devices may be divided into *descriptions* and *indexicals*. Reference by description occurs when the speaker is able to describe an object with precision adequate for the purposes to hand, without resorting to a proper name (of the object at the time of the utterance). The notorious example of Prince, a rock musician who unnamed himself and took on a symbol (different name?), is illustrative. Speakers quite successfully referred to him by (definite) description: *the artist formerly known as Prince*. Accepting for the moment

[11] Objects should be understood in a broad sense, to include events, processes, actions, social conventions, and so on, as well as physical objects such as shoes, ships, and sealing wax. Also, we do not say or require that the addressee 'understand' what the speaker is talking about, whatever that may mean. The question is whether the addressee can, operationally, identify the referent. See the article by Jones and Kimbrough in this volume for a discussion of signalling [JK04].

Russell's analysis [Rus05], "The artist formerly known as Prince is giving a concert in Philadelphia tomorrow" will be true just in case:

1. There exists an artist formerly known as Prince,
2. There is only one such artist, and
3. That artist is giving a concert in Philadelphia tomorrow.

(Further complicating the matter, I am told that as of the date of this writing, 29 August 2004, the fellow in question has reverted to calling himself Prince. Is our sentence still true? Well, suppose that Bobby Short, who has never changed his name, is playing at the Café Carlisle tonight. Is it true to say "The artist formerly known as Bobby Short is playing at the Café Carlisle tonight"?)

Besides describing things, the other main reference-fixing device, and perhaps a more basic one,[12] is simply to point to them, to indicate them in one way or another. This might be done non-verbally, say by pointing, nodding, or turning one's head. When it is done verbally, with language, we say that *indexicals* are used. Among them we will count proper names and pronouns. In addition, the *demonstratives*—in English, *this, that, these,* and *those*—are widely used in natural language. Here is an example from SeaSpeak, used for identifying the speaker over a public radio channel.

SeaSpeak Sentence 12 *This is Land's End Radio. [WGJS88, page 39]*

Interestingly, however, the SeaSpeak manual [WGJS88] does not have other *kinds* of examples using demonstratives. (The above form is used throughout to identify speakers.)

The use of *the*, as in SeaSpeak Sentences 5–8, remains unresolved. What are we to make of it? How are we to analyze it, given the linguistic and logical concepts described above?

Note that these SeaSpeak Sentences contain locutions that *resemble* definite descriptions:[13] *the dock entrance* (but there are many docks and each as at least one entrance), *the anchorage* (but there are many anchorages), *the icebreaker* (but there are many icebreakers), *the convoy*, and *the casualty.* Our suggestion is that reference gets fixed antecedently to these locutions, which may then be interpreted as definite descriptions *on the presupposition that the necessary reference-fixing has been done.* One might call this a *distributed definite description.* The suggestion is illustrated, and we think confirmed, by an example from the SeaSpeak manual [WGJS88, page 154]. See Figure 2.

[12] Since Peirce there has been a tradition of viewing demonstratives as the most basic form of reference [Hoo02], but that view has not gone unchallenged, e.g., [Kin01].

[13] The resemblance persists across multiple accounts of definite descriptions, including Russell's [Rus05], Strawson's [Str71], and Donnellan's [Don66] .

Kotka Radio	All ships in Gulf of Riga, all ships in Gulf of Riga. *This is* Kotka Radio, Kotka Radio. Ice information, area: Gulf of Riga. INFORMATION: The ice type is: winter fast ice, ice change: no change, ice navigation: ice-breaker assistance is necessary

Fig. 2. Distributed definite description example from SeaSpeak

The recommended SeaSpeak message structure is on display in Figure 2. Messages begin with header information and are followed by message elements, each of which begins with one of the message markers described above. Here in Figure 2 we see a header that makes three points:

1. "All ships in Gulf of Riga." This identifies the addressees. The message is being broadcast on an open radio channel.
2. "This is Kotka Radio." This identifies the speaker.
   Note: Identifying the addressee(s) and the speaker has to be done in the header of every message.
3. "Ice information, area: Gulf of Riga." This serves to specify what we call the *context of focus*. The speaker is announcing, declaring, that what follows is
   (a) Information pertaining to ice
   (b) Which ice is in the Gulf of Riga
   Note: Specifying the context of focus occurs in some but not all message headers (and messages).

The INFORMATION then consists of three assertions:

1. *The ice is (of type) winter fast ice.*
2. *The ice has not changed recently.*
3. *Ice-breaker assistance is required for those who wish to navigate the ice.*

In each case, we observe, *the ice* may be interpreted as *the ice in the Gulf of Riga*, producing the following stylistic variants of the three assertions:

1^1. *The ice in the Gulf of Riga is (of type) winter fast ice.*
2^1. *The ice in the Gulf of Riga has not changed recently.*
3^1. *Ice-breaker assistance is required for those who wish to navigate the ice in the Gulf of Riga.*

If *ice* were a singular count term (e.g., *the King of France*, *the man drinking a Martini*), then one's preferred analysis of definite descriptions would be appropriate (see below). In our example, however, *ice* is a mass term. Consequently, we propose the analysis evidenced in the following stylistic variants.

1^2. *There is ice in the Gulf of Riga and it is (of type) winter fast ice.*
2^2. *There is ice in the Gulf of Riga and it has not changed recently.*
3^2. *There is ice in the Gulf of Riga and ice-breaker assistance is required for those who wish to navigate it.*

We note that a stronger interpretation might be preferred:

1^3. *There is ice in the Gulf of Riga and all of it is (of type) winter fast ice.*
2^3. *There is ice in the Gulf of Riga and all of it has not changed recently.*
3^3. *There is ice in the Gulf of Riga and for all of it, if you wish to navigate it, then ice-breaker assistance is required.*

Other variants are possible. In what follows, unless otherwise noted, we employ variant 2.

7 Formalizing Distributed Descriptions into FLBC

Our principal goal in the present section is to combine the example from §5 with the analysis of *the*-expressions from §6. The result will be a structured, generalizable, and formal SeaSpeak message structure, akin to that in Figure 1, but incorporating distributed definite description.

7.1 Definite Descriptions and Nonexistence

Russell's analysis of definite descriptions is an excellent starting place for the method we shall employ here. Briefly, a definite description such as *The King of France is bald* would be analyzed as *There is exactly one thing that is King of France and it is bald.* Formally,

$$\exists x(F(x) \wedge B(x) \wedge \forall y(F(y) \rightarrow x = y)) \tag{1}$$

and similarly for other examples. Russell introduced useful notation for *the x such that* $\phi(x)$, viz., $(\imath x)\phi(x)$. Thus, Expression 1 may be abbreviated as

$$B((\imath x)F(x)) \tag{2}$$

This assumes an axiom schema permitting such formulations:[14]:

$$y = (\imath x)\phi(x) \leftrightarrow \forall x(\phi(x) \rightarrow y = x \wedge \phi(y)) \tag{3}$$

Russell's elegant theory has not gone uncriticized. From Expressions 2 and 3 it follows that $\exists x F(x)$, i.e., that something is King of France. This *might* be appropriate and perhaps in the present example it is. Perhaps, *The King of France is bald* is false just because (or sufficiently because) there is no King of France.

[14] Cf., [Lam91].

There cases in which we wish to speak, and think we can speak truly, about entities that do not or might not exist (in any straightforward way). *The surviving partner in Spade & Archer solved the murder* is—or can be taken as—true, even though we are talking about events in a work of fiction (*The Maltese Falcon* by Dashiel Hammett). If it is not true, then is it false? Such considerations lead naturally to the motivations underlying free logic, the variety of first-order logic in which names are free of existential commitment. Using the standard notation—$E!$—and definition of 'exists'

$$E!(y) \leftrightarrow \exists x(x = y) \tag{4}$$

the axiom schema in Expression 3 may be modified to be conditioned on the existence of the entity in question [Lam91]:

$$E!(y) \rightarrow y = (\imath x)\phi(x) \leftrightarrow \forall x(\phi(x) \rightarrow y = x \wedge \phi(y)) \tag{5}$$

Using expression 5 and given $B((\imath x)F(x))$ it follows that $E!(y) \rightarrow F(y)$, but in free logic, the inference from $F(y)$ to $\exists x F(x)$ is blocked absent $E!(y)$.

We shall steer a course permitting, but not requiring, free logic. We accept Expression 3 as an axiom schema or definition, but we require that in an formula of the form $y = (\imath x)\phi(x)$ that y be instantiated to a name, e.g., a_1. From $a_1 = (\imath x)\phi(x)$ and expression 3 it follows that $\phi(a_1)$ In standard logic this will entail $\exists x\phi(x)$ and in free logic this will entail only $E!(a_1) \rightarrow \exists x\phi(x)$. This affords us the flexibility of either using or discarding free logic as the case mandates.

When would a SeaSpeaker want to employ free logic and names for nonexistent things? It will often be useful to do so. Example: *The next arrival of the Rose Maru is scheduled for tomorrow.* Suppose that the Rose Maru fails to arrive. The sentence might still be true and if not perhaps the speaker is guilty of lying. Free logic will help get the inferences right and is convenient for this purpose.

7.2 Example

Figure 3, page 314, presents an FLBC representation of a SeaSpeak INFORMATION message with distributed definite description used for a *the*-expression. See Figure 2 for the *the*-expression in question. Figure 3 should be compared with Figure 1, page 306 in which a proper name is represented instead of a *the*-expression.

1. Reference-fixing
   s = the speaker
   a = the addressee
2. Presupposition
   (a) a_1 is the Gulf of Riga region:
   $a_1 = (\imath x)\mathit{Location}(x,$ *'Gulf of Riga'*$)$
   (b) a_1 is the focus of context for utterance e_1:
   $\mathit{Focus}(e_1, a_1)$
   (c) There is a state, s_1, of being ice at this time:
   $\mathit{ice}(s_1) \wedge\ \mathit{Hold}(s_1,\ \mathit{now})$
   (d) The state s_1 is located at a_1:
   $\mathit{Location}(s_1, a_1)$
3. Message body
   (a) Propositional content.
   There is ice in the Gulf of Riga and it is of type winter fast ice:
   $\exists x(\mathit{In}(x, s_1) \wedge\ \mathit{Type}(x,$ *winter-fast-ice*$))$
   (b) Illocutionary force.
   Further, the utterance is in the class INFORMATION, which indicates its illocutionary force. SeaSpeak FLBC Sentence 1 represents the INFORMATION force for the utterance to hand, with ϕ serving as a place-holder for the propositional content.
   SeaSpeak FLBC Sentence 4 $\mathit{information}(e_1) \wedge\ \mathit{Speaker}(e_1, s) \wedge\ \mathit{Addressee}(e_1, a) \wedge\ \mathit{Cul}(e_1,\ \mathit{now}) \wedge\ \mathit{Object}(e_1, \ulcorner\phi\urcorner)$

Fig. 3. FLBC representation of a SeaSpeak INFORMATION message with distributed definite description used for a *the*-expression

Points arising:

1. The substantive difference, between the message in Figure 3 (containing a *the*-expression) and the message in Figure 1 (with a proper noun instead of a *the*-expression), lies in section 2, holding the presuppositions.
2. Item (2a) of the message in Figure 3 uses a Russellian definite description to create a name—a_1—for the Gulf of Riga location, to be used as the context for the message.
3. Item (2b) declares that the contextual focus of the e_1 event (keying the *information* illocutionary force verb) is to be a_1, the Gulf of Riga *location* (as distinct from the Gulf of Riga itself).
4. Item (2c) is similar to item (2a) in Figure 1.
5. Item (2d) has no analog in Figure 1. It enforces the declaration of focus in item (2b).
6. With the presuppositions in order, the SeaSpeak FLBC Sentence is quite simple. In fact the illocutionary force components are identical in Figures 1 and 3.

8 Analysis of the Remaining SeaSpeak Message Markers

We have so far given two example analyses for SeaSpeak sentences with the INFORMATION message marker. It remains to discuss the other six message markers.

8.1 SeaSpeak WARNING Message Marker

In general we treat *warning* as an assertion with the presupposition that the speaker thinks there is related danger for the addressee.
Example sentence:

SeaSpeak Sentence 13 *WARNING: The Leading Lights are not lit.*

Stylistic variant:

SeaSpeak Sentence 14 *WARNING: The Leading Lights are unlit.*

Assumption: It is the Leading Lights of the ship Paisano that are not lit.

Figure 4 presents the FLBC message structure (with comments) for our WARNING example.

The points we wish to make about *the*-expressions have now been made. Consequently, for the remaining SeaSpeak message markers we will focus only on the treatment of the markers, eschewing treatment of the full messages, as in Figure 4. This will simplify and, we hope, clarify what follows.

8.2 SeaSpeak INTENTION Message Marker

Example sentence:

SeaSpeak Sentence 15 *INTENTION: I intend to reduce speed, new speed six knots.*

Under the analysis we offer, to say you intend that P amounts to asserting that you will see to it that P. Other analyses are possible and merit investigation, although the one on display is plausible, natural and, we think, quite serviceable. Our analysis, then, suggests the following stylistic variant:

SeaSpeak Sentence 16 *INTENTION: I assert that I will see to it that a speed reduction event occurs, with the new speed being six knots.*

Implicit in the context is that the speaker is referring to the speed of a particular vessel. Let us say it is the Paisano, a_1, as in Figure 4. Here, e_3 will be a *speed-reduction* event, whose *Theme* and *Focus* is a_1, the Paisano. What it is that is supposed to be seen to, then, is

1. Reference-fixing
   s = the speaker
   a = the addressee
   a_1 = *Paisano* (the registered ship of that name)
   LeadingLights(a_2) (a_2 has the property of being Leading Lights.)
2. Presupposition
   (a) The theme or topic of this message is a_2:
   Theme(s_1, a_2)
   (b) The focus of context for utterance s_1 is a_1:
   Focus(s_1, a_1)
   (c) The theme a_2 is located at a_1:
   Location(a_2, a_1)
   (d) The s_1 state is dangerous for a, the addressee:
   dangerous$(s_1) \wedge$ *Benefactive*(s_1, a)
3. Message body
   (a) Propositional content.
   a_2 (having the property of being Leading Lights located on the Paisano) is unlit:
   $(\mathit{unlit}(s_1) \wedge \mathit{In}(a_2, s_1))$
   (b) Illocutionary force.
   Further, the utterance is in the class WARNING, which indicates its illocutionary force is *assert*.
   SeaSpeak FLBC Sentence 5 *assert*$(e_1) \wedge$ *Speaker*$(e_1, s) \wedge$ *Addressee*$(e_1, a) \wedge$ *Cul*$(e_1,$ *now*$) \wedge$ *Object*$(e_1, \lceil$ $(\mathit{unlit}(s_1) \wedge \mathit{In}(a_2, s_1))$ $\rceil)$

Fig. 4. FLBC representation of a SeaSpeak WARNING message with a *the*-expression

SeaSpeak FLBC Sentence 6 *reduce-speed*$(e_3) \wedge$ *Theme*$(e_3, a_1) \wedge$ *Cul*$(e_3,$ *now*$)$

Seeing to it that is represented by the *stit* predicate, the FLBC analog of the *stit* operator in action logic.[15] The FLBC Stit Schema—

FLBC Schema 1 (FLBC Stit Schema) $\mathit{stit}(e_1) \wedge \mathit{Agent}(e_1, j) \wedge \mathit{Cul}(e_1, t_1) \wedge \mathit{Content}(e_1, \lceil \phi \rceil)$

—is rendered into English as *Agent j sees to it at time t_1 that ϕ.* Thus, adding the *stit* aspect to the speed reduction we get:

SeaSpeak FLBC Sentence 7 $\mathit{stit}(e_2) \wedge \mathit{Agent}(e_2, s) \wedge \mathit{Cul}(e_2, \mathit{now}) \wedge \mathit{Content}(e_2, \lceil \mathit{reduce\text{-}speed}(e_3) \wedge \mathit{Theme}(e_3, a_1) \wedge \mathit{Cul}(e_3,$ *now*$) \rceil)$

on the assumption that the speaker is s. Finally, we wrap all of this in an assertion to get:

[15] See discussion in [JK04], this volume, also [KY04].

SeaSpeak FLBC Sentence 8 $assert(e_1) \wedge\ Speaker(e_1, s) \wedge\ Addressee(e_1, a) \wedge\ Cul(e_1,\ now) \wedge\ Object(e_1, \lceil\ stit(e_2)\ \wedge\ Agent(e_2, s)\ \wedge\ Cul(e_2, now)\ \wedge\ Content(e_2, \lceil reduce\text{-}speed(e_3) \wedge\ Theme(e_3, a_1) \wedge\ Cul(e_3,\ now)\ \rceil)\ \rceil)$

8.3 SeaSpeak REQUEST Message Marker

The SeaSpeak manual [WGJS88, page 97] has this to say about requests:

> The word REQUEST will be used to signal messages which mean 'I want something to be arranged or provided' as in ships' stores requirements, bunkering, permission, ... It is commonly accompanied by the word **Please**, e.g.
>
> REQUEST: Please send, quantity: five acetylene cylinders.

SeaSpeak REQUESTs work, we think, more or less as ordinary requests as treated in the speech act literature. More specifically, a request differs from a command in that the request does not invoke any extraordinary authority. Recall the dialog in the movie *Casablanca*.

```
                LASZLO
Captain Renault, I am under your
authority.  Is it your order that we
come to your office?

                RENAULT
     (amiably)
Let us say that it is my request.
That is a much more pleasant word.

                LASZLO
Very well.
```

(See http://www.geocities.com/classicmoviescripts/ for a link to the script.)

Letting ϕ represent the propositional content of a REQUEST, we model the illocutionary force as follows:

SeaSpeak FLBC Sentence 9 $request(e_1) \wedge\ Speaker(e_1, s) \wedge\ Addressee(e_1, a) \wedge\ Cul(e_1,\ now) \wedge\ Object(e_1, \lceil\ \phi\ \rceil)$

We note that in the above example—*Please send, quantity: five acetylene cylinders*—it is not specified to whom the cylinders should be sent. Under the FLBC representation this would be indicated (or not) in the propositional content, ϕ, or in the presupposition. If, for example, the addressee is to send the material to the speaker, the content would include this fragment: $send(e_1) \wedge\ Agent(e_1, a)$ and the presupposition could include $Benefactive(e_1, s)$, perhaps by default.

8.4 SeaSpeak ADVICE Message Marker

The SeaSpeak manual [WGJS88, page 97] has this to say about the ADVICE marker:

> The word ADVICE will be used to signal suggestions, e.g.
>
> ADVICE: Anchor, position: bearing: one-nine-four degrees true, from Keel Point distance: one mile.

An advice, then, is a suggestion. We broadly agree with Searle and Vanderveken [SV85, page 202] that a suggestion in this sense is a kind of request (their category is *directive*). As we model them, suggestions (and SeaSpeak ADVICEs) are requests with the presumption that the beneficiary is the addressee. Thus, ADVICE comes with the presupposition that the event keying the propositional content, say e_2, benefits the addressee, a: $\mathit{benefits}(e_2, a)$. Note that *Benefactive* is a thematic role and does not imply anything good (or bad) for anyone. Even in ordinary English we can say "Brenda was the beneficiary of his tongue lashing and assorted insults" without suggesting there was anything good for Brenda in this. The special predicate $\mathit{benefits}(x, y)$ is to be interpreted as indicating a good of some kind, that y benefits (would benefit) from x occurring. Further, this is a general presupposition, unlike the referential examples above, in that it is not specific to the message to hand. Thus, it may be given as a general rule or axiom schema as part of the interchange agreement governing the communications. For development of this idea see [JK04] in this volume.

8.5 SeaSpeak INSTRUCTION Message Marker

The SeaSpeak manual [WGJS88, page 97] has this to say about the INSTRUCTION marker:

> The word INSTRUCTION will be used to signal commands, e.g.
>
> INSTRUCTION: Go to berth number: two-five.

We analyze commands as requests (or directives) in which the speaker invokes an authority and, if all goes well, puts the addressee under an obligation to honor the request. (See [JK04] in this volume for elaboration.) However, that there is an authority invoked and that the addressee is under an obligation need not be explicitly stated in the message. These are general conditions and they should be inferred from general rules governing the conversation. Consequently, our representation task here (and the burden on the communicants) is much reduced.

Even so, it will helpful for the speaker to declare the authority he is claiming. A simple and workable way to do this is to employ a special predicate,

AuthorityTitle(x, y, z),
with interpretation *x has and invokes the authority of y for the sake of z.* Thus, for example, *AuthorityTitle*(*s, 'Lighthouse Master',* e_1) added to the presupposition section would declare that speaker s is invoking the authority of a Lighthouse Master in uttering the messaged keyed by e_1. The illocutionary force of that message would have the form:

SeaSpeak FLBC Sentence 10 *command*$(e_1) \wedge$ *Speaker*$(e_1, s) \wedge$ *Addressee*$(e_1, a) \wedge$ *Cul*$(e_1,$ *now*$) \wedge$ *Object*$(e_1, \ulcorner \phi \urcorner)$

8.6 SeaSpeak QUESTION Message Marker

The SeaSpeak manual [WGJS88, page 97] has only this to say about the QUESTION marker:

> The word QUESTION will be used to signal all questions, e.g.
>
> | QUESTION: What is your ETA at the dock entrance? |
> |---|

Thus, the SeaSpeak QUESTION message marker is for the most part unproblematically interpreted as indicating a question in the usual, ordinary language sense.

We model questions as reference-fixing assertions recognized as presuppositions, plus requests to describe the referents. For example, a stylistic variant of "What time is it?" would be "There is some time that it is. [assertion presupposed] Please describe that time to me. [a request]" Similarly, "Senator, when did you stop beating your wife?" is a stylistic variant of "You used to beat your wife. Please describe when it was the beating stopped." Thus, a question (typically) makes a presupposition and goes bad if the presupposition is false or otherwise problematic. This is true of both w-questions (involving *who, what, where, when,* and *why*) and yes-no questions (answered with a 'yes' or a 'no'). We will examine one example of each kind.

W-Questions. Consider the what-question:

SeaSpeak Sentence 17 *QUESTION: What is your position?*

As usual, let s be the speaker and a the addressee. On our analysis, the speaker presumes that there is exactly one position that the addressee has, then requests that the addressee describe it. Thus, in the presupposition section of the message we would find

SeaSpeak FLBC Sentence 11 $x_1 = (\imath x)($*location*$(x) \wedge$ *At*$(a, x) \wedge$ *Hold*$(x,$ *now*$))$

which posits x_1 as a's one and only location. Additionally, s would state the presumption that x_1 exists: $E!(x_1)$.

The question is completed by a SeaSpeak (FLBC) sentence in which s requests that a describe x_1, a's one and only location.

SeaSpeak FLBC Sentence 12 $\mathit{request}(e_1) \wedge\ \mathit{Speaker}(e_1, s) \wedge\ \mathit{Addressee}(e_1, a) \wedge\ \mathit{Cul}(e_1,\ \mathit{now}) \wedge\ \mathit{Object}(e_1, \lceil\ \mathit{describe}(e_2) \wedge\ \mathit{Agent}(e_2, a) \wedge\ \mathit{Theme}(e_2, x_1) \wedge\ \mathit{Benefactive}(e_2, s) \wedge\ \mathit{Cul}(e_2,\ \mathit{now})\ \rceil)$

Yes-No Questions. Consider the yes-no question:

SeaSpeak Sentence 18 *QUESTION: Are you going to pass me soon?*

As a presupposition, the speaker expresses the event being asked about—

SeaSpeak FLBC Sentence 13 $\mathit{passing}(e_3) \wedge\ \mathit{Agent}(e_3, a) \wedge\ \mathit{Theme}(e_3, s) \wedge\ \mathit{Hold}(e_3,\ \mathit{soon})$

—a passing of s by a soon. The speaker does *not* presume that the event e_2 exists. That is the information requested in the SeaSpeak sentence:

SeaSpeak FLBC Sentence 14 $\mathit{request}(e_1) \wedge\ \mathit{Speaker}(e_1, s) \wedge\ \mathit{Addressee}(e_1, a) \wedge\ \mathit{Cul}(e_1,\ \mathit{now}) \wedge\ \mathit{Object}(e_1, \lceil\ \mathit{describe}(e_2) \wedge\ \mathit{Agent}(e_2, a) \wedge\ \mathit{Theme}(e_2, e_3) \wedge\ \mathit{Benefactive}(e_2, s) \wedge\ \mathit{Cul}(e_2,\mathit{now})\ \rceil)$

Because the speaker has not presumed $E!(e_3)$, i.e., that e_3 actually exists, the addressee is to infer by the conventions governing the conversation that a proper answer consists of asserting either that $E!(e_3)$ (the addressee will pass soon) or $\neg E!(e_3)$. Questions, we believe, inherently involve presuppositions, and when we presuppose about events and things that might not exist, free logic is a most handy tool.

9 Discussion and Conclusion

SeaSpeak, we noted at the outset, affords a natural experiment of sorts. It was conceived and developed without any apparent regard for doctrines of speech act theory or for suitability to formalization by an ACL (agent communication language). This and the fact that SeaSpeak is successful—it is used and entrenched in maritime communications—presents an apt challenge for particular ACLs and, more generally, the programme of formalization of business communication.

At least in a preliminary and theoretical way this challenge has been met by FLBC. First, we note that the message marker structure in SeaSpeak messages can only be seen as a heartening confirmation of a basic tenet of speech act theory, the $F(P)$ framework. Second, in two previous papers [KLPY03,KY04] we have explored with some success the translation

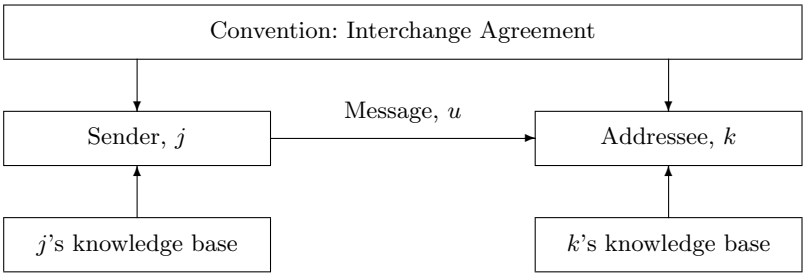

Fig. 5. Basic Messaging Framework: Message u from speaker j to addressee k (after [KT00])

of SeaSpeak sentences into FLBC, without at the same time addressing the problem of dynamic reference fixing. In the present paper, we have addressed this problem (at least in principle) and we have extended the analysis of SeaSpeak sentences to entire SeaSpeak messages. Significantly, these two issues are tied together by a common use of presupposition. In the case of *the*-expressions, we find a presupposition of a context of focus. In the case of questions, we find them to be inseparable from their presuppositions.

What about practical considerations? Is there any prospect in all of this for deployed applications? We think there is, but absent a working demonstration we opt not to stretch the reader's credulity (or our credibility). Nevertheless, we close with two observations. First, automated messaging occurs within an institutional context. See Figure 5 for a framework. That context includes conventions governing the exchange and interpretation of messages. Businesses using EDI even have a name for (a part of) this convention: the interchange agreement. It is the interchange agreement (or its equivalent) that determines, for example, how many message markers there are and what they mean. Our analysis shows that in SeaSpeak significant parts of a message belong to the reference-fixing or presupposition sections. These elements may often, then, be supplied by default, as we have remarked, because they are specified by a governing convention. In principle, what this means for practice is that with careful formalization *and* well-crafted conventions the information burden on the message composer can be greatly reduced. Messages may be 'lean' [KT00] and still convey all the essential information. Whether one is designing a user interface for humans or a messaging-generating program, this is welcome news.

Our second observation has to do with reference fixing. Many, if not most, of the problematic uses of *the*-expressions and related indexical reference-fixing devices can be handled as we have shown by fixing a context of focus and then relying on a (distributed) description. In ordinary language conversation this fixing of context is most often done by pointing in some way or by using a demonstrative expression (involving *this, that, these,* and *those*).

Is there a way for a human to instruct a machine in this way? How might a machine discern the object of an indexical by a human? In the simple case in which the context of focus has a proper name, e.g., the ship Paisano, there is little difficulty. The human interface is designed to capture the intended focus (say by picking from a list of names) and the computer program is given an established list of names and information about their objects. This simple case, we observe, may be generalized. Instead of mutually working with a common list of names, the communicants might be given maps or diagrams. A speaker might point (in any of various ways) to a region of a map and a supportive procedure could (for a suitably represented map) automatically extract the pointed-to context of focus. This might be done if the 'speaker' is a computer process and the addressee is a human, or vice versa.

These observations are, we think, straightforward enough. The consequences for applications can only be speculative at this point. We hope others will join us in exploring these possibilities.

References

[AEB04] Alan S. Abrahams, David M. Eyers, and Jean M. Bacon, *Practical contract storage, checking, and enforcement for business process automation*, Formal Modelling for Electronic Commerce: Representation, Inference, and Strategic Interaction (Steven O. Kimbrough and D. J. Wu, eds.), Springer, Berlin, Germany, 2004.

[Aus62] John L. Austin, *How to do things with words*, Oxford at the Clarendon Press, Oxford, England, 1962.

[Ben03] John Benyon, *Building police co-operation: The European construction site around the third pillar*, World Wide Web file, Accessed 28 February 2003, Dated 1996. `http://www.psa.ac.uk/cps/1996/beny.pdf`.

[BH79] Kent Bach and Robert M. Harnish, *Linguistic communication and speech acts*, The MIT Press, Cambridge, Massachusetts, 1979.

[BM85] David C. Blair and M. E. Maron, *An evaluation of retrieval effectiveness for a full-text document-retrieval system*, Communications of the ACM **28** (1985), no. 3, 289–299.

[Don66] Keith S. Donnellan, *Reference and definite descriptions*, The Philosophical Review **75** (1966), no. 3, 281–304.

[FBH86] Eileen Fitzpatrick, Joan Bachenko, and Don Hindle, *The status of telegraphic sublanguages*, Analyzing Language in Restricted Domains: Sublanguage Description and Processing (Ralph Grishman and Richard Kittredge, eds.), Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1986, pp. 39–51.

[Fos82] Antony Charles Foskett, *The subject approach to information*, 4th ed., Linnet Books, Hamden, CT, 1982.

[Har68] Zellig Harris, *Mathematical structures of language*, John Wiley & Sons, New York, NY, 1968.

[Har02] Richard K. Harrison, *Bibliography of planned languages (excluding Esperanto)*, World Wide Web, 1992-2002, accessed December 2002, `http://www.invisiblelighthouse.com/ langlab/bibliography.html`.

[Hoo02] Christopher Hookway, *Truth, rationality, and pragmatism: Themes from peirce*, Clarendon Press, Oxford, UK, 2002.

[J+93] Edward Johnson et al., *Policespeak: Police communications and language, English-French lexicon*, PoliceSpeak Publications, Cambridge Research Laboratories, 181a Huntingdon Road, Cambridge, CB3 ODJ, 1993, ISBN: 1 898211 01 9.

[Jac02] Ray Jackendoff, *Foundations of language: Brain, meaning, grammar, evolution*, Oxford University Press, Oxford, UK, 2002.

[JK04] Andrew J.I. Jones and Steven O. Kimbrough, *A note on modeling speech acts as signalling conventions*, Formal Modelling for Electronic Commerce: Representation, Inference, and Strategic Interaction (Steven O. Kimbrough and D. J. Wu, eds.), Springer, Berlin, Germany, 2004.

[Joh98] Edward Johnson, *LinguaNet final report*, World Wide Web file, 1998, http://www.hltcentral.org/usr_docs/project-source/linguanet/Final-Report/Final-Report.html, accessed 28 February 2003.

[Joh02] ———, *Talking across frontiers*, Proceedings of the International Conference on European Cross-Border Co-operation, 2002, Available at: http://www.prolingua.co.uk/talking.pdf. Accessed December 2002.

[Kim99] Steven O. Kimbrough, *Formal language for business communication: Sketch of a basic theory*, International Journal of Electronic Commerce **3** (Winter 1998–99), no. 2, 23–44.

[Kim90] ———, *On representation schemes for promising electronically*, Decision Support Systems **6** (1990), no. 2, 99–121.

[Kim01] ———, *Reasoning about the objects of attitudes and operators: Towards a disquotation theory for representation of propositional content*, Proceedings of ICAIL '01, International Conference on Artificial Intelligence and Law, 2001.

[Kim02] ———, *A note on the Good Samaritan paradox and the disquotation theory of propositional content*, Proceedings of ΔEON'02, Sixth International Workshop on Deontic Logic in Computer Science (John Horty and Andrew J.I. Jones, eds.), May 2002, pp. 139–148.

[Kin01] Jeffrey C. King, *Complex demonstratives: A quantificational account*, MIT Press, Cambridge, MA, 2001.

[KK01] Kyra Karmiloff and Annette Karmiloff-Smith, *Pathways to language: From fetus to adolescent*, Harvard University Press, Cambridge, MA, 2001.

[KLPY03] Steven O. Kimbrough, Thomas Y. Lee, Balaji Padmanabhan, and Yinghui Yang, *On original generation of structure in legal documents*, Proceedings of the International Conference on Artificial Intelligence and Law (ICAIL), 2003.

[KM97] Steven O. Kimbrough and Scott A. Moore, *On automated message processing in electronic commerce and work support systems: Speech act theory and expressive felicity*, ACM Transactions on Information Systems **15** (October 1997), no. 4, 321–367.

[KT00] Steven O. Kimbrough and Yao-Hua Tan, *On lean messaging with unfolding and unwrapping for electronic commerce*, International Journal of Electronic Commerce **5** (2000), no. 1, 83–108.

[KY04] Steven O. Kimbrough and Yinghui Yang, *Action at the tables: Sketching a tabular representation for utterances under the language–action perspective*, Proceedings of the 9th International Working Conference on the

Language Action Perspective on Communication Modelling (Rutgers, The State University of New Jersey, New Brunswick, NJ) (Mark Aakhus and Mikael Lind, eds.), School of Communication, Information, and Library Studies, 2–3 June 2004, http://www.scils.rutgers.edu/˜aakhus/lap/lap04.htm, pp. 103–120.

[Lam91] Karel Lambert (ed.), *Philosophical applications of free logic*, Oxford University Press, Oxford, U.K., 1991.

[LAP04] LAP, *Home LAP '04*, World Wide Web page, September 2004, http://www.scils.rutgers.edu/∼aakhus/lap/lap04.htm.

[Lar85] Andrew Large, *The artificial language movement*, Basil Blackwell, New York, NY, 1985, ISBN 0-631-14497-8.

[Lev83] Stephen C. Levinson, *Pragmatics*, Cambridge University Press, Cambridge, England, 1983.

[Lin03] LinguaNet, *The linguanet project*, 2003, www.cbs.dk/departments/fir/-linguanet/, accessed 2003-02-28.

[Mac30] T. C. Macaulay, *Interlanguage*, The Clarendon Press, Oxford, UK, 1930.

[Ogd38] C. K. Ogden, *Basic English : A general introduction with rules and grammar*, K. Paul, Trench, Trubner, London, UK, 1938.

[Pro03] Prolingua, *Operational communications, controlled languages, computing*, World Wide Web page, Accessed 28 February 2003, www.prolingua.co.uk/.

[Ric43] I. A. Richards, *Basic English and its uses*, W. W. Norton & Company, Inc., New York, NY, 1943.

[Rus05] Bertrand Russell, *On denoting*, Mind (1905), 479–93, Available at: http://www.mnstate.edu/gracyk/courses/web%20publishing/russell_on_-denoting.htm (accessed 30 July 2004).

[Sag75] Naomi Sager, *Sublanguage grammers [sic] in information processing*, Journal of the American Society for Information Science **26** (1975), no. 1, 10–16.

[Sag86] ———, *Sublanguage: Linguistic phenomenon, computational tool*, Analyzing Language in Restricted Domains: Sublanguage Description and Processing (Ralph Grishman and Richard Kittredge, eds.), Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ, 1986, pp. 1–17.

[Sea69] John R. Searle, *Speech acts*, Cambridge University Press, Cambridge, England, 1969.

[Str71] P. F. Strawson, *Logico-linguistic papers*, ch. On Referring, Methuen, 1971, Originally published 1957.

[SV85] John R. Searle and Daniel Vanderveken, *Foundations of illocutionary logic*, Cambridge University Press, Cambridge, England, 1985.

[Swa80] Charles Swanland, *Basic English for science & technology*, Intercultural Press, Chicago, IL, 1980.

[WGJS88] Fred Weeks, Alan Glover, Edward Johnson, and Peter Strevens, *Seaspeak training manual: Essential English for international maritime use*, Pergammon Press, Oxford, UK, 1988, Available via www.maritimeusa-.com.