

TR-IT-0008

Some Irregularities in English Conversation

Brian A. Yeager Yasuharu Den

1993.9

Abstract

This paper investigates four cases of unexpected word order found in conversational English, the Pseudo-cleft construction, uninverted sentences with question intonation, logophoric reflexive pronouns, and NPs that are not incorporated into standard syntactic structures. The discourse function of the various phenomena is considered as well their intonational characteristics.

ATR Interpreting Telecommunications Research Laboratories

©1993 ATR Interpreting Telecommunications Research Laboratories

Contents

1	Introduction	1
2	Data	1
3	Description of phenomena	2
3.1	Pseudo-cleft	2
3.2	Uninverted sentences with question intonation	5
3.3	Reflexives	8
3.4	Unincorporated NPs	12
4	Conclusion	12
	References	13

1 Introduction

This paper is concerned with four patterns that occur in English discourse that are likely to require special treatment in systems for automatic speech translation. They are the Pseudo-cleft, uninverted sentences with question intonation, logophoric reflexive pronouns and NPs which cannot be incorporated into a standard syntactic structure. There are many other phenomena relevant to the translation of natural English speech, and many of these are covered more extensively in the technical report written by Laurie Fais (Fais, 1993). Those interested in lexical discourse markers, interruptions, fragments and a much wider survey of the things that are likely to make free conversation difficult to process should look there. My objective was limited to identifying potentially problematic constructions frequently occurring in the sort of natural dialogue situations common in task-oriented cooperative dialogues.

Since the syntactic properties of these constructions are unclear, their functional role in English conversation will be investigated. For example, with the Pseudo-cleft the variations in word order may be off-set by a consistent discourse-marker function.

2 Data

I made primary use of the ATR International Conference Registration Database (ATR-CRD), which consists of 12,430 lines culled from telephone exchanges involving registration for several scientific conferences in addition to several calls to a travel agency regarding trips within Japan. The files are located in `atr-dp:/data11/ADD.TXT`, and they are split into the two sets 'Ryokou' and 'Kokusai', depending on their provenance.

I also made some use of the London-Lund Corpus (LLC) which has 122,655 lines. It is a collection of spontaneous, casual conversations with several participants with markings that convey prosodic information, and it is located at `as01:/export/London-Lund/londlund`.

The criteria by which the conversations were separated into lines was not made explicit, and so lines are probably not the best point of comparison between the two corpora. It seems that line breaks in the ATR-CRD reflect semantic and pragmatic utterances in some way. A sentence was never extended over a line break, and one speaker's turn might consist of several lines. The LLC, on the other hand, seems to be broken into lines by some type of phonological or prosodic criteria. A sentence could be broken across several lines with utterances from other participants intervening. When all the symbols conveying prosodic information were removed, the `wc` command¹ returned 497,508 as the number of words in the LLC corpus. This accords with the claim in the brochure that there are about 500,000 total words. There were approximately 120,536 words in the ATR-CRD by the same criteria. This is a better measure of their relative sizes, although discrepancies in the treatment of phenomena such as of hesitation noises and unintelligible segments still remain.

Although these sort of variations in transcription style make comparisons difficult, the basic trend of differences between the two corpora was relatively clear. The free conversation recorded in the LLC featured longer turns, more interruptions and less-salient discourse

¹For `wc` a word is a string delimited by the SPACE, TAB or NEWLINE characters.

objectives than the clearly task-oriented conversations of the ATR-CRD. Although the goal for speech translation technology may be closer to the LLC style of spontaneous speech, there were several notable patterns to be found in the ATR-CRD that can be expected to cause problems for machine translation.

3 Description of phenomena

3.1 Pseudo-cleft

The most striking recurrent example of an ungrammatical form occurring in the International Conference Registration Telephone Dialogue Corpus were the utterances that mutated the Pseudo-cleft construction. In what follows, I will examine how these sorts of utterances were used and how I think they can best be analyzed in a Machine Translation system. Since cleft constructions have been justly recognized as focus based, the analysis will necessarily touch on several important issues in processing natural utterances.

There are many instances of utterances with unexpected word order involving what appears to be the Pseudo-cleft construction. When I looked into it more carefully, I found that the ungrammatical utterances associated with this construction are closely circumscribed. Of approximately 27 utterances featuring the Pseudo-cleft construction, only about half are ungrammatical, and almost all of those fall into two groups.

Unexpected word orders arise when the head verb of the preposed clause is one of a set of verbs consisting of propositional attitude verbs and verbs of communication that could be called cognitive verbs (*think of, mean, ask, know*) and when the head verb is *do*.

As far as the cognitive verbs are concerned, one often (about 6 times out of 7) sees expressions like the following in the transcripts.

Kokusai 95

What I'd like to ask you is: could you pay me, could you make that sum in cash after a tax deduction?

correct: What I'd like to ask you is whether you could pay me ...

What is taken to be the complement clause of *ask* is inverted. This, however, is not a problem specific to the Pseudo-cleft construction, since embedded clauses are often inverted in other contexts in English as well. For example, someone might informally report a conversation in the following terms.

John asked could Mary come.

The empirical facts aren't clear as far as I know (proscriptively, inverting is prohibited, but common sense says this is an extremely common pattern), but probably these verbs will need to have two subcategorization frames, one for an embedded uninverted clause and one for an inverted clause. This should carry over to the Pseudo-cleft in any reasonable grammar.

The Pseudo-cleft constructions with the verb *do* in the preposed clause are somewhat less-easily handled. There are a total of 17 such utterances in the two corpora, with 11 in the ATR-CDR and 6 in the LLC. Of those, three utterances with *do* feature complement clauses with infinitival clauses, but the most common pattern is represented by the 11 sentences that are irregular in the following way.

Kokusai 102

What I will do is I will send you the slides in three packets.

correct: What I will do is send you the slides in three packets.

Instead of the bare VP that is expected as the complement to the dummy verb *do*, many speakers produced utterances with a full sentence after the copula. There are several options for writing a grammar that will be able to handle constructions like this. The simplest is that you could simply specify two different subcategorization frames for the verb *do*. On the other hand, there is some motivation for analyzing these utterances as two different syntactic structures: a Pseudo-cleft with a null deictic predicate followed by a sentence explicating the null pronoun. There are several points in favor of this analysis. Colloquial expressions like the following seem to exhibit the same phenomenon without *do*.²

The thing is he can't see.

An overt deictic pronoun is in fact possible, as in the following utterance.

What I'll do is this. I'll call you.

and the fact that this strategy would also work for the mystery inversions involved with the cognitive verbs mentioned above. One problem with this idea is that there are (colloquial) utterances like *I'll call you is what I'll do* that seem less amenable to such an analysis.

In the end the decisive criterion in deciding whether such utterances consist of one or two syntactic structures will probably be their prosodic characteristics. Differences in timing and accent may indicate whether these utterances are best treated as involving two utterances, and even if not, it seems likely that these utterances have a characteristic intonation contour that will identify them. It seems to me that there is necessarily strong stress on the *do* in such cases. This hypothesis was not disproved by the LLC. Of the six examples of the pseudo-cleft found there, three displayed a distinctive intonation pattern.

LLC 02

b: darling I am sorry about that champagne

a: fifteen

²Constructions of the form *The thing...do* is occurred three times in the LLC. Twice the predicate was an infinitival VP and once a full clause beginning with *for*. Depending on whether these are considered possible antecedents for null pronouns, the comparison with the Pseudo-cleft may or may not be appropriate.

that's a point
you see
^what you d/o is
you go ptshoing
Gordon

Two of the three utterances showing this pattern featured full sentential complements. Of the three unaccented sentences, only one featured a full sentential complement, and the Pseudo-cleft construction itself was embedded in that utterance, so the meager data available seem consistent with the hypothesis that there is a distinctive prosodic signature associated with Pseudo-clefts with full sentential complements, even if they do not fully confirm it.³

A characterization of the function of the Pseudo-cleft construction would be helpful, but even though it has been commonly acknowledged to have a topicalization function, it is not clear from the corpus exactly what this means, if it is true. Consider a typical example.

Kokusai 102

Q: What type of trays do you use?

Do you use [um,] the Kodak carousel style or do you use box-stack style trays?

[Um,] so I know how to set these things up before I come.

S: [Uh,] as for the multi-slides are concerned, (you) you can use the multi-slide equipments, and you can use the remote controls.

And the tray we are gonna use is the carousel.

And [uh,] we will set the carousel on the slides.

However, we have to confirm the order of the slides before (you) your presentations.

And there is a slide reception at the entrance of the conference hall.

Therefore, you will confirm the order of your slides (and) at the (sli) slide reception, and slide operator will set the slides on the equipments.

Q: I understand.

What I will do is I will send you the slides in three separate packets since I'll be using three projectors.

³There were 24 instances of the Pseudo-cleft in 120,536 words of the ATR-CDR and only 6 in the 538,879 words of the LLC. I think that this difference reflects the difference in the types of conversations recorded in the two corpora. If one accepts that the Pseudo-cleft is conventionally used as a kind of discourse-segment concluding discourse marker (see below), then its frequency in task-oriented telephone dialogues and scarcity in free conversation is unsurprising. In the sort of natural speech recorded in the LLC, there is little need to explicitly structure one's utterances.

They'll be numbered [um,] A, B, C, A1 <A one>, A2 <A two>, A3 <A three>, B1 <B one>, B2 <B two>, B-, so there really shouldn't be any problem.

Here the Questioner is a lecturer who is discussing the logistics of how his presentation will proceed with a member of the planning section. He asks several questions (in addition to the one noted above) about how the slides will be presented, and after hearing the response from the Secretariat, he responds using the Pseudo-cleft. In broad strokes, the usage given above typifies those found in the corpus in that the Questioner is employing the Pseudo-cleft construction to indicate that he is summarizing his plans in light of new information from the Secretariat. I think that the best way to approach describing the function of the Pseudo-cleft object construction is to think in terms of discourse markers, as discussed in Grosz and Sidner (1986), but the model described there has difficulties with real speech. There doesn't seem to be any difference in the function of correct and incorrect Pseudo-clefts in discourse.

3.2 Uninverted sentences with question intonation

There are a variety of utterances in the transcript with question intonation that do not have either auxiliary inversion or a wh-word at the beginning. This characterization actually describes many disparate phenomena. Some of them, such as the analysis of fragmentary and confirmational utterances, are bound to be important in any attempt to analyze real human speech, so a more detailed treatment is in order.

Any analysis of phenomena associated with question intonation (which I am assuming is uniformly represented by a question mark in the text, even though this is a fairly tenuous assumption in my opinion), will necessarily involve a characterization of what a question is, and this in turn involves one with speech acts and illocutionary forces. In my preliminary look at these matters, I've been operating with a binary opposition between statements that are confirmations (epitomized by expressions like "you mean NP" where NP is a repetition of previous discourse) and information questions, more or less in concert with what I take to be the predominant sentence labelling practice around here. Still, these labels bear scrutiny, since there are often cases where a hearer announces a conclusion he has arrived at, or confirms the basic content of what has been said without employing the same terms. In these situations and others like them it's not clear whether these two illocutionary force labels suffice.

The first type of utterance, actual Wh-in situ, really doesn't occur very often. I don't think that it occurs intentionally, so I don't think that it carries any information.

When wh words are actually present, examples of Wh-in situ are easily recognizable (though there remains a problem as to how or even whether to provide for them in an MT system). When there is no wh word, that is, when the utterance is a case of an uninverted yes-no question, there is the possibility for significant confusion between what I've called an information question (when an answer yes or no is sought) and a confirmational utterance. An example of the first sort of utterance is the following.

Kokusai 154

S: Well, we will send the copies of the proceeding and a program to you.

So upon the receipt of such materials could you kindly transfer the money through the bank account or by check?

Q: Well, [ah,] (maybe I could just) yeah, OK.

I'll send you a check.

That would be fine.

[Um,] A personal check [ah,] would be fine?

S: Well in principle we are regretting [ah,] to take the personal checks, but [ah,] isn't there any company check available?

An example of the second sort would be something like this.

Kokusai 83

S: And at the other day we have received the abstract of you and we typed it already.

And [uh,] we send it the last week it to you.

And have you received it?

Q: Yes, (I) I got it.

[Uh,] there're also [ah,] a few problems with that, (I'm) I imagine that you've seen.

I'd like to make a few corrections and make a few [ah,] spelling corrections and a few [ah,] corrections to the text.

Do you have the copy in front of you?

S: Yes, I have here.

But [ah, eh,] in case, [ah,] now you are going to tell us what is the problem here.

But [ah,] we fear that we will make [ah,] another misspelling or something.

And we have already [ah,] we have two weeks before the conference will start, [ah,] therefore, [um,] maybe you can send it to you then by the post then we will make the correction.

Q: You're going to send me a copy and I'll correct it and send it back to you?

(You) you don't ...

(How) how much time do you need before the conference to reproduce it?

I'll be glad to do what you've asked but I wanna make sure we have enough time.

A human looking at these two examples in context feels a difference, but the status of the difference is not clear. There seem to be some indications that the speakers' intentions differ in the two cases, since usually when a speaker used a confirming expression, he continued on with his discourse turn, assuming what he'd said to be correct whereas there were answers to inquiries like that in Kokusai 154. On the other hand, if the speaker in there had used the most natural yes/no question available, he would have said *Would a personal check be OK*. The difference in degree between *OK* and *fine* may be reflected in the choice of conversational acts. When one wants to know whether it is acceptable to use a check, one uses the neutral value *OK*, but if one is indicating something by assuming something, one might assume the maximal possible degree.

This use of uninverted sentences with question intonation as a confirmation was the primary pattern for native speakers of English. Non-native speakers used a lot more uninverted utterances as real informational questions.

Perhaps the most significant aspect of this phenomenon is that it is not limited to full sentences. Utterances consisting only of an NP, for example, with question intonation were very common. As in the following example, fragments with question intonation were usually, but not always, confirmations.

Kokusai 46

Q: (then, then) then you want (a complete) a complete version later?

Is that right?

(If) if my summary ...

S: Just a moment.

Q: [Huh, huh.]

S: Hello.

Q: Hello.

[Uh,] you don't need to submit the complete version.

Q: [Huh, huh.]

S: Just summary is all right.

Q: Just the summary?

S: Just summary is enough, yes.

Q: Just the summary.

[Um,] I'm a little unclear.

(You) you said [uh,] submission of copies of 200 <two hundred> words summary.

Finally, it should be noted that uninverted sentences with question intonation carry different presuppositions than utterances with tag questions. There were no instances of the uninverted sentence with contrastive focus as well as question intonation in the corpus.

3.3 Reflexives

I did a search for the strings *self* and *selves* in the ATR-CRD and for the string *self* in the LLC. There were two results that I think are interesting from an MT perspective. The first is that there are many cases of reflexive pronouns occurring in positions where other NPs are not allowed, and the second is that there was a significant split in the sorts of reflexives found in the two corpora.

The search of the ATR-CRD yielded a total of approximately 72 utterances containing reflexives, and of these there were 26 instances of *myself* (34%), 17 instances of *itself* (24%), 15 instances of *yourself* (20%), and various others. The first and second person reflexives could occur in what are called logophoric positions, that is positions where NPs are not usually found, and they occurred there quite frequently.

Kokusai 88

I unfortunately don't know much about it, having never left Michigan myself.

The third person reflexives occurred quite frequently as appositives in what you might be tempted to call focusing constructions.

Kokusai 58

Is there a full [ah,] translation service available for the conference itself?

These two patterns alone account for 43 of the 72 instances of reflexives found in the ATR corpus. Additionally, there were many instances of prepositional phrases of the type *by ... self* and *for ... self*. (The string *by myself* was especially common.) Whatever the eventual analysis of these prepositional phrases is, it is clear that the reflexives occurring in them are not the canonical sort of reflexive that attaches a reflexive meaning to a predicate. If we include the reflexives occurring in such prepositional phrases with the appositives and free reflexives noted above, then 54 of the 72 (75%) reflexives uttered in the ATR-CRD are logophoric, without antecedents in their clauses.

Comparing this pattern with the distribution of reflexive pronouns found in the LLC is illuminating. There there were only 11 logophoric reflexives among a total of approximately 260 unaccented reflexives. That is, among the lines returned by searching for the strings *self* or *selves*, there were very few logophoric pronouns. There were no occurrences of *by ... self*, and only one occurrence of *itself* as an appositive. The vast majority of reflexive pronouns occur in what might seem their more typical usage, a sentential anaphor reflexivizing the predicate which it occurs with.

LLC 04

b: and he keeps coming

to self-defence

a: (laughs)

- b: you'd think he'd be old enough to either keep
 out of trouble
 or ((defend his))
- a: ((all this sitting around on street corners
 with his bicycle chain 2 sylls))
- c: (laughs)
- b: yeah
- b: or defend himself

However this striking distribution pattern only obtains for the unaccented reflexives in the LLC database. When the search was extended to include accented reflexives, the disparity disappeared. With the accented reflexives included there were a total of 507 instances of reflexive pronouns in the LLC, giving a total of approximately .1% of words uttered being reflexives as compared to the total of approximately .6% in the case of the ATR-CRD. There were 24 instances of *itself* and many instances of *by ... self* and *for ... self*. Among the accented reflexives, a large majority were logophoric. A conservative survey of 80 clear cases revealed that 60 were logophoric. Since this survey excluded lines with nothing but bare reflexive pronouns, and reflexive pronouns are likely to be logophoric in isolated utterances, it seems likely that the actual percentage is higher, and perhaps even in the same range found in the ATR-CRD. Prototypical examples of the two sorts of logophoric pronouns follow.

LLC 04

- a: I've never been to Harlow
 is it nasty
- b: well
 the actual building's all right
 it's a great ((sort of building)) like a modern
 ivory tower
 you know
 which is rather refreshing
 if you're working in a place where everything's
 falling to bits like
- a: yeah
 yeah

b: that part's all right
but H\arlow its\elf⁴
is a great windswept new town

LLC 04

a: I think I'll go to ((sleep d'you))
b: yes
I feel a bit sleepy mys\elf

The data in the LLC make clear that reflexive pronouns do not bifurcate into logophoric and anaphoric on the basis of whether or not they are accented. There are accented anaphoric reflexives and unaccented logophors. Still, there is a clear tendency for the prosodic status of a reflexive to reflect its syntactic status, since an unaccented reflexive was a logophor less than 5% of the time.

This prosodic tendency was matched by a functional consistency. In the ATR-CRD reflexives were primarily used for two specific purposes, both of which fit under the general rubric of focusing techniques. When plans were being made in which one party would not assist the other, the *by ... self* construction was often used. This function seems to be specific to this sort of planning dialogue. For instance, the following sentence.

Kokusai 41

S: [Uh,] I'm so sorry, but we cannot arrange any reservation for the hotel.
And call-for paper has [uh,] some informations about the hotel accommodations.
Q: [Huh, huh.]
S: Would you please refer to that and make reservation **by yourself**?
Q: I see.
So I have to do that all **by myself**.
S: [Huh, huh.]

The other pattern is best characterized by saying that when the speaker wanted to emphasize a discourse referent, a bare reflexive—an appositive or a floating pronoun displaced to an unexpected position such as the end of the sentence or immediately before or after the verb—was used. This pattern was less homogeneous. The appositive *NP itself* pattern almost always occurred as *the conference itself* which seems to indicate that lately the focus of the conversation has been something related to the conference but that now the speaker wants to talk about only the conference. The floating reflexive pattern seems similar.

⁴The prosodic symbols have all been removed from the LLC examples for ease in reading, except for examples like this where they are crucial. The \ symbol probably represents what is labelled a L accent in other systems.

Kokusai 44

Q: (First) first of all, I was only planning to audit the conference.

S: [Huh, huh.]

Q: [Uh,] but I've noticed I have to submit a 200 < two hundred > summary.

S: [Huh, huh.]

Q: Now, (wh, what sort) what sort of summary do you mean?

I mean I have no plans to contribute anything myself.

In sentences like these, it's clear that any informational content that the reflexive has is at an extremely high level, along the line of connectives like *but* or some types of intonation contours. In this typical example, the reflexive seems to explicitly contrast the Questioner from other people. "Other people might contribute something but, as for me, I have no plans to contribute anything". The Questioner thinks that the Secretariat thinks that the Questioner plans to submit something, and with this utterance he wants to assert that of all the propositions of the type that they are talking about, that is of all the propositions of the form *X plans to contribute something*, the instantiation of the proposition with Q is not true. The effect, then, seems to be equivalent to the open sentence that Rooth (1985), Pierrehumbert and Hirschberg (1990) and others have postulated for utterances with constituents receiving focus stress. This utterance does seem to be very similar to the following sentence where the subject receives focus stress.

I have no plans to contribute anything.

It seems plausible that logophoric reflexives consistently receive prominent stress. If this were the case then one could approach the difference between this last, focus-stressed sentence and the previous example with a logophoric reflexive as a sort of *iinaosi*, in which information that should have been conveyed by the prosodic contour of the base utterance is made up for by including an extra word at the end of the utterance. The speaker wanted to focus the subject in the sentence but didn't when it was actually pronounced, so he had to make up for it later in the utterance. This is especially plausible for utterances involving first and second person subjects, since putting focal stress on first or second person subjects may be dispreferred for stylistic reasons. Without an investigation of the intonational characteristics of this sort of reflexive it is impossible to make any authoritative statement, but the data seem to suggest that a system taking constituents' accentual status into account may be able to analyze logophoric reflexives handily.

Logophoric reflexives may give some clues about the general phenomenon of how fragments with no clear connection to other syntactic structure present in an utterance should be treated. While not as common in English as in Japanese there were occasional examples of free NPs uttered after a sentence had been completed.

3.4 Unincorporated NPs

I don't have too much to say about this—it wasn't a terribly common phenomena in the corpus I looked at. This seems to be a phenomena correlated with the register of a conversation, in that it occurs relatively more frequently in the entirely conversational LLC corpus more often than in the conference registration corpus. Laurie Fais noted several occurrences of this sort of utterance, and I saw one.

Kokusai 48

... are they expensive, the taxis here ...

This seems to me to be connected to the problem of fragments in general, and fragments with question intonation in particular. From an MT perspective the big question seems to be what to do with this sort of NP once you've determined that they don't belong to any utterance. As this example suggests, they may have some bearing on pronoun construal, so perhaps they need to be kept track of. Incidentally, they occur in patterns reminiscent of the syntactic topicalization construction.

Sapir and Whorf, they were linguists.

This sentence is naturally used in at least two contexts with accents on *Whorf* and *they* respectively.

Finally, it might be noted that since NP floating is not a syntactic phenomenon, unincorporated NPs can be related to any position in a sentence, including positions inside Wh, relative clause, and adverbial clause islands.

John ate it, the cookie.

John asked what they would eat, the visitors from Spain.

John left before they came, the policemen.

John saw the present they had left, the welcoming committee.

4 Conclusion

That prosodic information is important in analyzing speech should surprise no one. What is especially tantalizing about these corpora is that both of them are based on actual recordings, and so prosodic information really should have been available. Although it is so disparate that domain specific patterns will be difficult to detect, I think that the LLC will be helpful for this reason. Also, since for the time being it seems that speech translation efforts may be confined to specific domains like registering for international conferences, it will be helpful to know how these sort of conversations vary quantitatively from unconstrained, multi-participant, casual conversations.

The results of this research suggest that there may indeed be qualitative differences in the types of constructions and the prosodic patterns used in different sorts of discourse situations. We have seen that the Pseudo-cleft occurs much less frequently in spontaneous conversation than in task-oriented discourse, that more often than not it features an unexpected word order pattern, and that this pattern may be correlated with a characteristic prosodic signature. Question intonation was found in utterances with a variety of intentional forces. The majority of reflexive pronouns were logophors with no antecedent in the same clause, and there was a clear tendency for unstressed reflexives to function in the more typical role of anaphors. Finally, similar to logophoric reflexives, several NPs were found in positions where they are not expected, and since they carried information useful in pronoun construal, a high-quality speech recognition system should make use of them.

References

- [1] Fais, L. "Non-grammatical Phenomena in Real English Conversation," *ATR Technical Report TR-IT-0007*, 1993.
- [2] Grosz, B. and C. Sidner "Attentions, Intentions, and the Structure of Discourse," *Computational Linguistics*, 12 (1986), 175-204.
- [3] Pierrehumbert, J. and J. Hirschberg "The Meaning of Intonational Contours in the Interpretation of Discourse," in *Intentions and Communication in Discourse*, MIT Press 1990.
- [4] Rooth, M. *Association with Focus*, University of Massachusetts Dissertation, 1985.
- [5] Walker, M., M. Iida and S. Cote "Centering in Japanese Discourse," in *Proc. of the 13th International Conference on Computational Linguistics*, 1990.