FORMS OF INTRODUCTION IN MAP TASK DIALOGUES: CASE OF L2 RUSSIAN SPEAKERS

Olga V. Goubanova

Centre for Speech Technology Research, University of Edinburgh, 2 Buccleuch Place, Edinburgh EH8 9LW, UK
olga@cstr.ed.ac.uk

ABSTRACT

In the paper forms of introduction (question or non-question) of new information into a discourse by American English and L2 Russian speakers are studied. 48 Map Task dialogues between native and L2 speakers are analyzed. Forms of introduction are found to be influenced by speaker role, order of a map presentation, and native language of a speaker factors. The choice of a particular form of introduction is also influenced by whether some of the landmarks on instruction giver’s, follower’s, or both participants’ maps were lacking labels. Communicative success of a conversation is assessed as deviation scores of an instruction follower’s route from a giver’s route. Order of a map presentation is found to affect deviation scores. When a map was introduced for the first time, a route drawn by an instruction follower was less accurate than the one s/he drew for the second time. Furthermore, a route drawn by a L2 instruction follower was less accurate than the one reproduced by a native follower. In addition, factors influencing length of a conversation are studied. L2 speaker proficiency level is found to affect length of a conversation. Conversation length is also affected by difficulty of the task; when a map has some landmarks with missing labels (label altered condition), it took longer for the speakers to complete the task.

1. INTRODUCTION

For a conversation to proceed smoothly, speakers should understand each other’s contributions, especially when new information is introduced into a discourse. It is known that non-native speakers differ from native speakers in the way they communicate in second language (L2). Our initial hypothesis is that L2 speakers should use different forms of introduction of new information into a discourse. We chose Map Task as a research tool to test this hypothesis.

Over the past 15 years, Map Task has been used extensively in studying such linguistic phenomena as variability of speech [1], conversational structure and intonation [2]-[3], and forms of introduction in dialogues [4]. In the work by [4] question and non-question forms of introduction, with each type of article, used by Scottish speakers of English were studied. The goal of the present research is to look at forms of introduction used by L2 Russian speakers in the corpus of Map Task dialogues [5] between native and non-native English speakers. To assess communicative success of the dialogues, we measured deviation scores of an instruction follower’s route from a giver’s route for each dialogue in the corpus. We also investigated what factors influence length of a conversation. The structure of the paper is as follows. We talk about the method used to study forms of introduction in section 2. We discuss the experimental results in section 3. We make conclusions and talk about future work in section 4.

2. METHOD

2.1. Materials

For the purpose of the present research we used a corpus of Map Task dialogues between native and L2 Russian speakers [5]. (Throughout the paper, we refer to these dialogues as L2R Map Task corpus.)

To remind, Map Task involves two participants. One, an instruction giver (IG), has a map with a route on it. The other, an instruction follower (IF), doesn’t have a route on her/his map. An IG describes a route in relation to the landmarks s/he has on her/his map; an IF draws a route on her/his map as closely as possible to the IG’s map. In the L2R Map Task corpus 3 out of 16 pairs of the HCRC Map Task Corpus [6] maps were used. From this set of 3 pairs of maps additional 9 pairs of maps were created by removing some of the landmarks’ labels on an instruction giver’s, follower’s, or both participants’ maps. Therefore, the L2R Map Task dialogues were collected under 4 experimental conditions; neutral (no landmarks’ labels on either IG or IF maps were removed), some of the landmarks’ labels on an instruction giver’s map were removed (“IG map altered” condition), some of the landmarks’ labels on an instruction follower’s map were removed (“IF map altered” condition), and labels were removed on both participants’ maps (“IG+IF maps altered” condition). Approximately 30% (3 − 4) of the map’s landmarks were altered, i.e., they were lacking labels, so that a subject has to describe a landmark using a referring expression of his/her own. In the L2R corpus, the HCRC experimental design in terms of the number of dialogues each subject can produce (2 as an instruction giver with the same map, and 2 as an instruction follower but with two different maps) was preserved. Therefore, a sextuple of speakers (3 pairs of native-L2) produced 12 dialogues per condition, with the whole corpus constituting 48 dialogues. For the present analysis we selected all 48 dialogues elicited from 12 American English and 12 L2 Russian speakers.

2.2. Procedure

We classified an introduction of a landmark as question (Q), if an utterance was in interrogative form marked by syntax or intonation (example 1). The introductions that were embedded in a route description were classified as non-question (NQ) (example 2).
the listener’s knowledge regarding the same landmark. We con-
sidered these as NQ introductions (example 3). Some landmarks,
that were lacking labels due to the experimental conditions, were
introduced using referring expressions. We still classified such a
landmark introduction as Q or NQ given a referent was understood
from the dialogue context (examples 4, 5). We did not analyze
forms of introduction based on definite-indefinite articles due to the
confounding effect of Russian L2 speakers’ errors (examples 2, 6).
1. And do you have walled city?
2. And after that you go straight. *You see a Canadian Paradise.
3. Now there is a walled city. Do you have city?
4. I just have eh something that looks like the map of the United
States without a name on it.
5. And do you have something underneath the walled city?
6. *Do you have picket fences?

3. RESULTS AND DISCUSSION
3.1. Forms of introduction analysis
We calculated the number of introductions for question and non-
question categories for each of the 48 dialogues. We performed
an analysis of variance with the number of introductions as the de-
pendent variable. The within-conversation variables were speaker
role \textit{ROLE} (instruction giver or follower), native language of a
speaker \textit{LANG} (English or L2), and form of introduction \textit{FORM}
(Q or NQ). The between-conversation variables were order of a map
presentation \textit{ORDER} (first or second time), and label altered con-
dition \textit{COND} (neutral, "IG map altered", "IF map altered", or
"IG+IF maps altered"). As was expected from the analysis by [4],
the ANOVA showed a significant effect of speaker role \textit{ROLE}
(\textit{F}(1,128) = 39.44; \textit{p} < 0.001); on average instruction givers
introduced more landmarks per dialogue than instruction follow-
ers (8.27 vs. 3.94). There were a number of interactions found.
There was one double interaction between form of introduction and
speaker role \textit{FORM ⋆ ROLE} \textit{FORM}(1,128) = 11.284; \textit{p} < 0.001). On average, instruction givers used more Q and NQ introductions
than instruction followers. The mean number of Q and NQ intro-
ductions per dialogue were 5.06 and 3.21 per instruction giver, and
1.65 and 2.29 per instruction follower, respectively. Therefore, as the task was becoming more challeng-
ing, instruction givers adopted a more cautious style of introducing
new information into a discourse. Instead, instruction followers
presupposed more shared knowledge with their partners about map
landmarks and used a more terse style of landmark introduction.

There were two triple interactions found as well. One was
between form, speaker native language, and map order \textit{FORM ⋆ LANG ⋆ ORDER}(\textit{F}(1,128) = 4.301; \textit{p} < 0.05). The results of
mean number of introductions per dialogue by form by map order
used by native and L2 speakers are shown in Figure 1. As can be
seen from the figure, when giving a map for the first time, both L2
and native speakers used more Q than NQ form introductions, with
the mean number of Q and NQ introductions per dialogue being
3.08 and 2.88 per L2 speaker, and 3.79 and 2.67 per native speaker,
respectively. When attempting their second map task, L2 speakers
maintained the same cautious style of introducing a landmark; they
introduced it using more Q than NQ form introductions (3.75 vs.
2.08). Conversely, native speakers reversed their behavior during
the second attempt of the map task; they used more NQ than Q
introductions (3.13 vs. 2.75).

Another triple interaction found was between form of introduc-
tion, label altered condition, and speaker role \textit{FORM ⋆ COND ⋆ ROLE} \textit{FORM}(3,128) = 5.516; \textit{p} < 0.001). The results of mean num-
ber of introductions per dialogue by form by label altered condition
used by instruction givers and followers are shown in Figure 2. As
can be seen from the figure, instruction givers tended to use more
question than non-question form introductions compared to neutral
condition, with the mean number of Q form introductions per dia-
logue being 3.92, 6.17, 6.17, and 4.0 for neutral, "IG map altered",
"IF map altered", and "IG+IF maps altered" conditions, respect-
ively. Under the same experimental conditions instruction follow-
ers tended to use more NQ than Q form introductions compared to
neutral condition, with the mean number of Q form introductions
per dialogue being 1.58, 2.5, 2.0, and 2.58 for neutral, "IG map altered",
"IF map altered", and "IG+IF maps altered" conditions, respec-
tively. Therefore, as the task was becoming more challeng-
ing, instruction givers adopted a more cautious style of introducing
new information into a discourse. Instead, instruction followers
presupposed more shared knowledge with their partners about map
landmarks and used a more terse style of landmark introduction.

There was also one quadruple interaction found, that one being
between form, language, map order, and speaker role \textit{FORM ⋆ LANG ⋆ ORDER ⋆ ROLE}(\textit{F}(3,128) = 5.325; \textit{p} < 0.025). In
Figure 3 the results of mean number of introductions per dialogue
by form by map order used by native and L2 givers/followers are
shown. As can be seen from the figure, when giving a map for
the first time, L2 givers used as many questions as non-questions to
introduce a landmark; the mean number of Q and NQ introductions
per dialogue being 4.33 and 4.0. On the second attempt, Russian
givers used over 2 times more Q than NQ form introductions, thus
adopting a more cautious communication style. Conversely, na-
tive instruction givers used almost 2.5 times more questions than
non-questions on the first attempt of a map. When giving the same
map again, they used almost as many Q as NQ form introductions.
Russian instruction followers used slightly more Q than NQ in-
tructions when receiving a map both times. Native followers,

![Fig. 1. Mean number of introductions per dialogue by form by map order used by native and L2 speakers (FORM ⋆ LANG ⋆ ORDER).](image-url)
instead, used more Q than NQ form introductions. (See a table from Figure 3).

3.2. Measuring communicative success

As a measure of communicative success we calculated a deviation of an instruction follower’s route from a giver’s route using centimeter squared grid, as was suggested in [4]. In addition, the ANOVA on the route deviation data with map order, instruction giver native language, and label altered condition variables confirmed the main effect of map order ($F_{(1,31)} = 4.177; p < 0.05$). Figure 4 shows the results of mean deviation scores by landmark altered condition by map order for native and L2 instruction givers. As can be seen from the figure, under all experimental conditions the map deviation scores are higher on the first attempt of a map presentation for both native and L2 givers. Moreover, the deviation scores are higher for native than for Russian instruction givers on the first as well on the second attempt of a map task. As can be seen from the figure’s table, on the first attempt of a map presentation, the mean deviation scores for the dialogues with native instruction givers range from 101.17 to 164.17 cm$^2$; for the dialogues with Russian L2 givers the scores range from 78.67 to 145.25 cm$^2$. It may well be that Russian instruction followers were having comprehension problem when given instructions by native instruction givers, whereas native instruction followers understood L2 givers’ directions without a problem.

3.3. Factors influencing length of a conversation

In addition, we studied factors that may have an effect on length of a task-oriented conversation. We assumed that in a task-oriented dialogue between a native and a L2 speaker, length of a conversation should be influenced by a number of factors such as native language of an instruction giver, label altered condition, and order of a map presentation. We performed an analysis of variance, with length of a conversation being the dependent variable; and native language of an instruction giver $GLANG$, label altered condition $COND$, and map order $ORDER$ being the independent variables. The main effect of label altered condition $COND$ was significant ($F_{(3,47)} = 4.125; p < 0.05$). No interactions were significant.

3.3. Factors influencing length of a conversation

In addition, we studied factors that may have an effect on length of a task-oriented conversation. We assumed that in a task-oriented dialogue between a native and a L2 speaker, length of a conversation should be influenced by a number of factors such as native language of an instruction giver, label altered condition, and order of a map presentation. We performed an analysis of variance, with length of a conversation being the dependent variable; and native language of an instruction giver $GLANG$, label altered condition $COND$, and map order $ORDER$ being the independent variables. The main effect of label altered condition $COND$ was significant ($F_{(3,47)} = 4.125; p < 0.05$). No interactions were significant.

Fig. 2. Mean number of introductions per dialogue by form by label altered condition used by instruction giver/follower ($FORM \times COND \times ROLE$).

Fig. 3. Mean number of introductions per dialogue by form by map order used by native and L2 givers/followers ($FORM \times LANG \times ORDER \times ROLE$).

Fig. 4. Mean deviation scores per dialogue by map order by label altered condition for native and L2 instruction givers.

We also considered language proficiency of a L2 speaker $PROF$ as a factor affecting length of a conversation. In a post hoc analysis we divided all the L2 participants into three groups; good, fair, and bad speakers. The good speakers group (4 speakers) consisted of the L2 speakers who had good verbal command of the L2 language; they could communicate in English with almost no pronunciation or grammar errors. The fair speakers group (5) communicated satisfactorily with their native partners; they made just a few errors in pronunciation and grammar. The bad speakers group (3) consisted of the L2 speakers who could not converse freely in English; they made a lot of pronunciation errors and grammar errors. We performed the ANOVA on the durational data with proficiency and label altered condition variables as independent variables. The main effect of $COND$ was significant ($F_{(3,47)} = 7.576; p < 0.001$). The interaction between label altered condition $COND$ and L2 speaker proficiency level $PROF$ was found to be significant as well ($F_{(3,47)} = 8.179; p < 0.001$). Therefore, conversation length is affected by L2 speaker proficiency. It is also influenced by difficulty of the task at hand, which
is driven by label altered condition variable.

4. CONCLUSIONS AND FUTURE WORK

From the analysis of the data it follows that speaker role affects her/his choice of forms of introductions, with instructions givers using more introductions than instructions followers. Furthermore, form of introduction is affected by native language of a speaker and order of a map presentation. When presenting a map for the first time, both L2 and native speakers used more question than non-question form introductions. On the second attempt, Russian speakers still used more Q than NQ form introductions, maintaining the same cautious communication style. Conversely, native speakers reversed their behavior during the second attempt of the map; they used more NQ than Q introductions.

When the task had become more challenging, instruction givers adopted a more cautious style of introducing new information into a discourse; they used more Q than NQ form introductions. Instead, instruction followers presupposed more shared knowledge with their partners about map landmarks and used more NQ than Q form introductions.

Order of a map presentation was found to affect accuracy of a map produced by an instruction follower. When an instruction giver introduced a map for the first time, a route drawn by an instruction follower was less accurate than the one drawn by another instruction follower, when s/he was receiving instructions from the same giver. Furthermore, the conversations between a native giver and a L2 instruction follower resulted in a less accurate route drawn. Therefore, the high deviation scores for the conversations with Russian instruction followers can be indicative of the communication difficulties non-native followers were experiencing.

It was found that L2 speaker proficiency affects length of a conversation; the conversation between a native and a bad L2 speakers took longer than either the one between a native and a fair speakers, or the one between a native and a good L2 speakers. Conversation length is also affected by difficulty of the task; when a map has some landmarks with missing labels (label altered condition), it took longer for the speakers to complete the task.

In this paper we used the methods of classical statistics to discover dependency relations among the data. In the future, we plan to apply Bayesian approach to modelling forms of introduction used by L2 and native speakers, similar to the one described in [7]. A set of Bayesian variables will consist of discrete variables such as form of introduction type (question or non-question), map order (first or second), native language of a speaker (English or L2), speaker role (instruction giver or follower), label altered condition (neutral, "IG map altered", "IF map altered", or "IG+IF maps altered"), with the number of introductions being a continuous Gaussian node. Given such a network, we would be able to perform inference on the conversation data. For example, we could calculate the probability of a person being a non-native speaker of English based on the values of other variables in the network. Such an analysis could be potentially used for accent and/or speaker identification tasks. In addition, this inferential information may have application in new/given information disambiguation in discourse structure analysis.

5. ACKNOWLEDGEMENTS

The author would like to thank Bob Ladd and Ellen Gurman Bard for the valuable discussion of the L2R Map Task corpus design. The author is also grateful to Stephanie Shattuck-Hafnagel for arranging for the L2R Map Task recordings at Speech Communication Laboratory, MIT.

6. REFERENCES