Dialogue strategies for... hur säger man ordförrådsträning?

Andrea Carrión Del Fresno and **Staffan Larsson** and **Vladislav Maraev**

Dept. of Philosophy, Linguistics
and Theory of Science
University of Gothenburg
`guscarrian@student.gu.se,`
`staffan.larsson@ling.gu.se,`
`vladislav.maraev@gu.se`

Abstract

We carry out a small-scale empirical study of a dialogue strategy (conversational pattern) found in second language learner dialogues where a language-assisting teacher is present, allowing learners to pick up new words and train on them while maintaining a conversation. We also provide a formal model of the observed conversational pattern including several frequently occurring variants, as well as a demonstration implementation which is able to reproduce the most common variant of the pattern.

1 Introduction and previous work

We are interested in dialogue strategies for vocabulary training in second language learner's dialogues. By finding and analysing recurring patterns in human-human dialogues, we hope to provide a solid empirical basis for the implementation of dialogue strategies in dialogue systems for second language learning.

This paper combines a small-scale empirical study of a conversational pattern similar to that described in Svennevig (2018) in second language learner dialogues. We also provide a simple computational model and a demonstration implementation which is able to reproduce the most common variant of the pattern. We believe such an implementation can be a very useful addition to conversational systems for second language learners. In Section 2, we describe the corpora and tools used, and then move on to the corpus study in Section 3. The formal model resulting from the corpus study is presented in Section 4, and the implementation based on the formal model is explained in Section 5. In Section 6, we provide conclusions and in Section 7 we describe future work.

2 Resources

The dialogue excerpts used in this paper were extracted from two different second language learner spoken corpora, namely, the European Science Foundation Second Language Databank (ESF)[1] and the Barcelona English Language Corpus (BELC)[2]. Both corpora belong to the SLABank collection[3], a part of TalkBank responsible for providing corpora in order to study the field of second language acquisition and learning.

The ESF database collects spontaneous conversations between adults of different nationalities that are learning a second language, including Dutch, English, German and Swedish, and native speakers of those languages. It should be noted that only those conversations where English is the target language were used in this study.

The BELC corpus collects speech recordings of Spanish students between the ages of 8 and 18 who are learning English as a second or even third language (Catalan is also spoken in the area where the research was conducted).

Importantly, there is also an investigator present, providing language support when needed. The investigator interacts with the subject using the target language, although it is shown that the investigator also knows the subject's mother tongue and resorts to it if necessary.

2.1 TalkBank browser and SCoRE

In order to access and collect the data, both the TalkBank Browser and SCoRE were used. The former is a browsable database that lets you navigate through transcripts from various corpora as well

---

[1] `https://slabank.talkbank.org/access/Multiple/ESF/`
[2] `https://slabank.talkbank.org/access/English/BELC.html`
[3] `https://slabank.talkbank.org/access/`

as watching or listening to any audio or video files attached to them, if available.

SCoRE[4] is a tool for browsing dialogue corpora, originally intended to search the British National Corpus (BNC) but now also able to access other corpora. The web interface allows the user to easily search a corpus with the help of regular expressions.

3 Data collection

In the data collection phase, we collected dialogue excerpts where a production problem together with code-switching take place.

A number of search expressions were used to make the search for proper dialogue examples more efficient. We used some sentences in both the target language (English) and the first languages (Italian, Spanish, Catalan) spoken by the subjects in the cited corpora that could serve to identify those potential examples: "how do you say", "come si dice", "cómo se dice", "com es diu". In addition, we searched for clarification ellipses, i.e. turns that repeated a word from the previous turn and that were understood as questions

The search process resulted in a collection of 40 suitable dialogue extracts.

4 Data annotation

The target dialogue extracts were manually annotated using a taxonomy of dialogue acts that was created for the purpose of this work but based on previous related taxonomies (Varonis and Gass, 1985; Bondarenko, 2019; Howes et al., 2019; Myrendal, 2019)[5]. Table 1 below shows a detailed description of the annotation tags that make up the taxonomy. We use the following abbreviations:

- S, S1, S2: speaker
- INV: Investigator (teacher)
- SUB: Subject (learner)
- L1: learner's first language
- L2: target language that learner is acquiring
- M: word or phrase in L2 that learner is missing

A sample of 10 dialogue transcripts as well as a description of the annotation tags and some in-

[5]The dialogue act taxonomy used here makes a number of fine-grained distinctions that are beyond the scope of more general dialogue act annotation schemas like DAMSL (Core and Allen, 1997) or the ISO standard (Bunt et al., 2017).
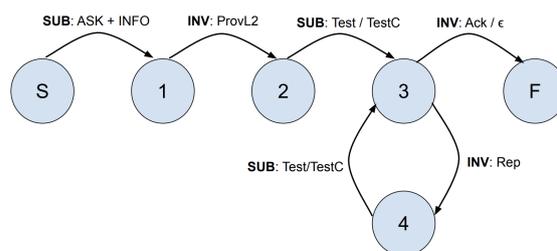

Figure 1: Final-state diagram representing a recurring pattern found in the data.

structions were provided to two annotators in order to ensure inter-rater reliability. Fleiss' kappa test showed a score of 0.812 which indicates good agreement. An example of an annotated dialogue excerpt is shown in Table 2.

5 Formal model

Based on the annotated corpus of dialogue excerpts, we analysed dialogue act sequences looking for recurring patterns with the goal of providing a simple formal model, preferably in the form of a finite state automaton. We found that while a wide variety of dialogue act sequences were used to initiate the repair sequences, they thereafter largely followed a predictable pattern with some minor variations. We therefore split the formal model into two phases where the initial phase (Ask+Info) is separated out from the overall model.

As seen in Figure 2 the formal model presents a finite state automaton with a total of six states, with S and F being the initial and final state respectively. Each action performed by both the subject and the investigator represents the transition from one state to the next one.

State S to 1: The transition between the initial state and state 1 corresponds to the initial 'Ask + Info' phase, which includes some way of asking for a translation of a missing word, about which some information is provided (typically, it's L1 form). We will describe this phase further in Section 5.1. Lines 104 and 105 in the example in 2 above provide an example of behavior in this transition. Specifically, the subject is unable to find the word "traghetto" in English (i.e. "journey"), and consequently he/she asks "what's the name?" of the word in the target language.

State 1 to 2: In this transition the missing information is provided by the investigator as Table 2 shows in line 106 in 2.

Tag	Description	Example utterances
AskL1	SUB asks for word using L1	S: ¿Cómo se llama? ⟨What is the name?⟩ S: ¿Cómo es? ⟨How is it?⟩
AskL2	SUB asks for a word using L2	S: What is the name?
IndAskL1	SUB indirectly asks for a word using L1	S: no sé el nombre ⟨I don't know the name⟩ S: no sé cómo se dice ⟨I don't know how to say it⟩
IndAskL2	SUB indirectly asks for a word using L2	S: I don't know what the name is. S: I don't remember how to say the word.
SearchL2	SUB (unsuccessfully) searches for a word or phrase in L2	S: The price of food is... eer... is... S: I only read books and... er... S: We bought tomatoes and... mm...
ProvL1	SUB provides the L1 translation of M	S: ...mesa ⟨table⟩ S: ...¿niña? ⟨girl?⟩
ProvL2	INV provides M in L2	S: It's called a table.
ProvDesL2	SUB describes M in L2	S: This thing you use for brushing your hair
Rep	INV repeats M	L: Mobile phone T: Mobile phone ←
Test	SUB tries pronouncing M	T: It's called a sprinkling can. L: Sprinkling can. ←
TestC	SUB tries using M in context	T: It's called a sprinkling can. L: We took the sprinkling can to water the plants. ←
Ack	S2 acknowledges previous utterance by S1	S1: We went to the park on Friday. S2: mhm

Table 1: Dialogue act annotation schema

Line	Spkr	Text	Annotation	Stage
104	SUB	yeah... and during the.... traghetto ⟨journey⟩	SearchL2 + ProvL1	1
105		what's the name?	AskL2	1
106	INV	journey.	ProvL2	2
107	SUB	journey?	Test	2
108	INV	journey.	Rep	2
109	SUB	journey.	Test	2
110	INV	mm.	Ack	2
111	SUB	during the journey.	TestC	2

Table 2: Annotation of dialogue liean24i.1.cha

State 2 to 3: The subject repeats the information (line 107 in 2) given in the previous state as a way to (a) practice the correct pronunciation and (b) reinforce the acquired knowledge and/or even (c) let the investigator know that the conversation can now proceed.

States 3 to 4 & 4 to 3: These transitions are an optional repetition of M by both the investigator and the subject. This behaviour can take place once or several times, as long as the investigator considers that this repetition is necessary for the subject's proper acquisition of the new information before the conversation can continue. Lines 108 and 109 in 2 illustrate the optional transitions.

State 3 to F: Once the subject successfully repeats the new information, the investigator accepts the subject's contribution to the conversation (line 110 in 2). This can be done overtly using a verbal

acknowledgement, or silently[6]. At this point the main conversation is ready to resume (line 111 in 2).

5.1 First stage of the model, Ask+Info

Here, we will not focus on describing patterns in the 'Ask+Info' stage of the model, that is, dialogue act combinations observed in our dialogue extracts during the transition between the initial state and state 1. See (Carrión del Fresno, 2022) for a description of this stage.

5.2 Second stage of the model

The most repeated structure is the sequence ProvL2 + Test + Ack, present in 40% of the data.

> S: and straight on in the ⟨pause⟩ street er ⟨pause⟩ the ⟨pause⟩ come si dice la strada principale ⟨*how do you say the main road*⟩.
> I: the main road. [ProvL2]
> S: mh the main road. [Test]
> I: mh. [Ack]

Slightly different to this pattern is ProvL2 + TestC + Ack (at 20%), where the subject is testing the new information in context. That is, the subject does not just repeat the provided information but uses it to continue the conversation:

---

[6]Because we are lacking video recordings of the interactions, we do not know if acknowledgement was provided gesturally, e.g. using a head nod.

S: mm ⟨pause⟩ ⟨pause⟩ ma mi scorde sempre come si dice la porta ⟨*i always forget how to say door*⟩.
I: door. [ProvL2]
S: mm door ⟨pause⟩ ⟨pause⟩ er ⟨pause⟩ ⟨pause⟩ no open. [TestC]
I: yeah. [Ack]

6   Implementation

The idea behind the implementation was to reproduce a dialogue strategy frequently observed in our data and embed it in a vocabulary training activity in the second language classroom. The dialogue focuses on vocabulary related to food where the main topic of the conversation revolves around what the learner has had for breakfast. By (verbally) interacting with the dialogue system, the learner is able to reinforce the acquired knowledge of the target language but also learn new lexical items.

For our implementation,[7] we used statecharts (Harel, 1987) which allows to describe the complex behavior of a system using an extended finite state notation. In addition, we chose to work with XState[8] for the model implementation. It is a JavaScript library designed to interpret finite state machines and statecharts in a way defined by Harel and W3C SCXML standard (Barnett et al., 2015).

The fact that we deal with code-switching in our dialogue excerpts makes it crucial to rely on a bilingual ASR (Automatic Speech Recognition) so that the dialogue system can handle a conversation where the learner alternates between both the first and second language. This is made feasible by setting two individual ASRs with separate confidence scores, corresponding to the native and the target language. In our implementation, Spanish and English are taken as the user's first and target language respectively.

An example dialogue with the system could go as follows:

S: What did you have for breakfast?
U: I had toast with... cómo se dice queso?
S: Cheese.
U: Cheese?
S: Uh-huh. Did you have anything else?

We see that the implemented model as a blueprint of a mechanism that be used in the language classroom for practising new words in the context of simulated everyday practical conversations such as making reservations, buying travel tickets, checking in at a hotel, etc..

7   Conclusions and future work

The main goal of the current study was to investigate dialogue strategies for vocabulary learning that could be found in second language learner corpora, and that could be useful in a dialogue system for second language training. The formal model encapsulates a general strategy used among learners at the time of acquiring new vocabulary in the second language, when in the presence of a teacher who can offer language assistance.

The proposed model is based on data from students of English with a poor linguistic competence where the described production problems were common. The vocabulary building activity is designed for learners of a second language at early stages where linguistic support is often needed due to the lack of knowledge in the target language. However, whilst production problems may not be as frequent among advanced learners, the strategy described seems to be still applicable to any learner regardless of their level of linguistic competence.

Future work includes extending the implementation to cover more variants of the patterns observed in the repair initiation (phase 1). We would also like to explore larger quantities of data provide an even stronger empirical footing. Also, conducting a human evaluation within the second language learning context would be a key component in future attempts to evaluate the model's performance. We would also like to confirm the applicability of the model to other language pairs, and in particular involving second languages other than English.

Acknowledgements

---

[7] https://github.com/guscarrian/breakfast_demo
[8] https://xstate.js.org

References

J Barnett, R Akolkar, RJ Auburn, M Bodell, D Burnett, J Carter, S McGlashan, T Lager, M Helbing, R Hosn, et al. 2015. State chart XML (SCXML): State machine notation for control abstraction, W3C recommendation.

Anastasia Bondarenko. 2019. *Grounding of names in directory enquiries dialogue. A corpus study of listener feedback behaviour*. MA thesis, Master in Language Technology, Gothenburg University.

Harry Bunt, Volha Petukhova, David Traum, and Jan Alexandersson. 2017. Dialogue act annotation with the ISO 24617-2 standard. In *Multimodal interaction with W3C standards*, pages 109–135. Springer.

Andrea Carrión del Fresno. 2022. Dialogue strategies for vocabulary training: User initiative in dialogue systems for second language learning. *M.A. thesis, University of Gothenburg*.

Mark G Core and James Allen. 1997. Coding dialogs with the DAMSL annotation scheme. In *AAAI fall symposium on communicative action in humans and machines*, volume 56, pages 28–35. Boston, MA.

David Harel. 1987. Statecharts: A visual formalism for complex systems. *Sci. Comput. Program.*, 8(3):231–274.

Christine Howes, Anastasia Bondarenko, and Staffan Larsson. 2019. Good call! grounding in a directory enquiries corpus. In *Proceedings of the 23rd Workshop on the Semantics and Pragmatics of Dialogue-Full Papers*.

Jenny Myrendal. 2019. Negotiating meanings online: Disagreements about word meaning in discussion forum communication. *Discourse Studies*, 21(3):317–339.

Jan Svennevig. 2018. "what's it called in norwegian?" acquiring l2 vocabulary items in the workplace. *Journal of Pragmatics*, 126:68–77.

Evangeline Marlos Varonis and Susan Gass. 1985. Non-native/non-native conversations: A model for negotiation of meaning. *Applied linguistics*, 6(1):71–90.