Edinburgh Research Explorer

I wanna talk like you

I wanna talk like you: Speaker Adaptation to Dialogue Style in L2 Practice Conversation

Arabella J. Sinclair[1], Rafael Ferreira[1], Dragan Gašević[2,1], Christopher G. Lucas[1], and Adam Lopez[1]

[1] University of Edinburgh, Informatics Forum, 10 Crichton St, Edinburgh, EH8 9AB, United Kingdom {s0934062,dragan.gasevic, alopez, cglucas2}@ac.ed.uk
[2] Monash University, 19 Ancora Imparo Way, Clayton, VIC, 3800, Australia
dragan.gasevic@monash.edu

Abstract. We present a novel method for analysing speaker alignment in second language conversational practice dialogue. In particular, we represent utterances as Dialogue Acts, and use Epistemic Network Analysis to analyse their usage patterns. We hypothesise convergence between speakers will be visible at this level, and influenced by learner ability and tutor strategy. We find as student ability increases, the distribution of student and tutor Dialogue Acts is both initially closer, and converges more over the course of the dialogue. We also find that the type of Dialogue Act use changes both with ability and throughout an interaction. Our results contribute a novel method for the automatic analysis of both learner ability and tutor strategy, which can inform both the development of personalised automatic tutoring tools, and formative assessment and feedback in an educational setting.

Keywords: Epistemic Network Analysis, Scaffolding, Dialogue, Natural Language Processing, Alignment, Zone of Proximal Development

1 Introduction

One-to-one spontaneous dialogue practice represents an important aspect of Second Language (L2) learning in both a classroom setting and online learning platforms. This form of dialogue practice has been shown to provide greater opportunities for L2 learning [7, 1, 18, 2, 9] as learners can both take advantage of the example of their interlocutor, and learn through practice. We investigate two main aspects of alignment: (i) symmetry between speakers within dialogue, and (ii) convergence of speakers over the course of an interaction, exploring effects of learner ability and speaker role across these aspects. We use Dialogue Acts (DAs) [8] to label utterance roles and analyse alignment at this level. DAs are used to infer discourse structure, and for automatic understanding of spontaneous dialogue [22]. DAs provide a high level, topic-agnostic method for speaker contribution comparison. DAs are particularly suited to the analysis of L2 learning, where the language itself is the educational content. We thus arrive at the following research questions: RQ1: *What is the relationship between students' and tutors' DA usage and student ability?* and RQ2: *How does the distribution of DAs usage change over the course of a dialogue?* We hypothesise that DA usage will be more similar as student ability improves, as speaker contributions become increasingly symmetric [4] and that speakers will converge within a single dialogue [21]. We employ Epistemic Network Analysis (ENA) [19] to model how DAs co-occur within student-tutor one to one dialogues at different levels of student ability, quantifying speakers' dialogic contribution.

Our main contributions are twofold: Firstly, we contribute to the existing literature on speaker adaptation within L2 dialogue, providing evidence to support our hypothesis that alignment can be seen at the level of DAs both with increasing ability level and across dialogues. Second, we provide a novel method for modelling both students' and tutors' contributions within conversational dialogue practice through combining the descriptive powers of ENA with DAs. This has implications for both formative assessment in an instructional setting, and continuous feedback for tutors and students, providing a data-informed reflection on their practice. Our work also has implications for (i) the design of learning analytic tools, (ii) informing tutoring strategy, and (iii) the design of automatic tutoring systems.

2   Background

Alignment consists of interlocutor interaction adaptation, resulting in convergence, or in their sharing of the same concept space [14, 3, 6]. Typically, alignment is measured at either a *lexical* (use of the same words [26, 20] or phrases [5] as each other) or a *syntactic* (use of the same parts of speech patterns e.g. similar noun-phrase constructions, or similar adjuncts [15, 16]) level. Methods for measuring alignment range from count statistics [5] to linear regression on prime-to-target distance[3] [25] to using general linear mixed models to account for the random effects present in dialogue [16, 20].

Dialogic interaction is said to play a central role in all learning [10]. Alignment in dialogue has been used as a predictor of both student learning and engagement [25], and it has been hypothesised that learners may leverage alignment to improve achievement of pedagogic goals [12]. Within an L2 context, students also show effective learning through interaction [7], although the potential for alignment is shaped by different speaker goals [4]: L2 learners benefit from vocabulary and grammatical examples provided by their interlocutors, but learners' ability will affect how useful those examples are. L2 learners also perform at a higher level when speaking with a peer than in a monologue context [17], suggesting they draw from the language of their interlocutor, leading us to expect some evidence of alignment. The *tutor's* goals also affect alignment: Vygotsky's Zone of Proximal Development (ZPD) [24] states students learn best when addressed at the correct level, thus we expect tutor adaptation to student ability. Viewing dialogue as a collaborative learning process, we expect to see speakers approach a shared understanding at the utterance exchange level [9]. While we expect speakers to arrive at a communicative symmetry [23], the nature of a tutor-student relationship has an expert-novice asymmetry. That said, at higher student ability, a tutor will alter their role to that of conversational peer to better encourage student independence and autonomy, thus slowly removing some support [2]. This change in tutor role is one aspect of ZPD which we examine at the level of labeled interaction sequences via *DAs*, dialogue labels describing utterance function, such as *question, statement* or *backchannel*[4].

3   Data & Methods

The dataset used is the Barcelona English Language Corpus (BELC) [13]. It consists of 118 transcripts from conversational practice between English learners and tutors. These vary from 60 to 140 utterances in length. Tutors' instructions were to elicit as much naturalistic conversation from the learner as possible, setting students at ease. A similar topic script was followed for all dialogues. The corpus is divided into four general levels of student ability ranging from absolute beginner to intermediate. We use DA annotations from our previous work [21], chosen from [22] for their relevance to the corpus. Table 1 shows the DA labels used and provides example dialogues from the highest and lowest level students, which demonstrate the differences in DA use between student and tutor.

We model the relationship between speakers' DA distribution and student ability via Epistemic Network Analysis (ENA) [19]. ENA is a graph-based analysis method for examining associations between different concepts (*codes*) in textual datasets. Two codes are considered related if they appear in the same *stanza*, which in our case are either the full dialogue (RQ1), or dialogue quarters divided by number of utterances (RQ2). ENA produces graphs that capture the relationships between different codes within an *analysis unit*, an individual or a group of participants. In our work, each DA is a code, and speakers at each level a unit of analysis. Each utterance is represented as a vector of the presence (1) or absence (0) of each code. A co-occurrence matrix is then derived for each dialogue from these code vectors. Dimensionality reduction is then performed using Singular Value Decomposition (SVD) [11] providing the visualisation of the projection graph in a two dimensional space i.e. [svd1, svd2]. The code relationships for the units of analysis are represented as undirected graphs, the size of the node representing the frequency of each code, and the edge width the strength of co-occurrence. ENA also allows us to explore the trajectory of a unit of analysis over time, within the same projection graph.

---

[3] The item being aligned to in this context is known as the *prime*, and the subsequent usage of this prime by the other speaker is known as the *target*, or sign of alignment.
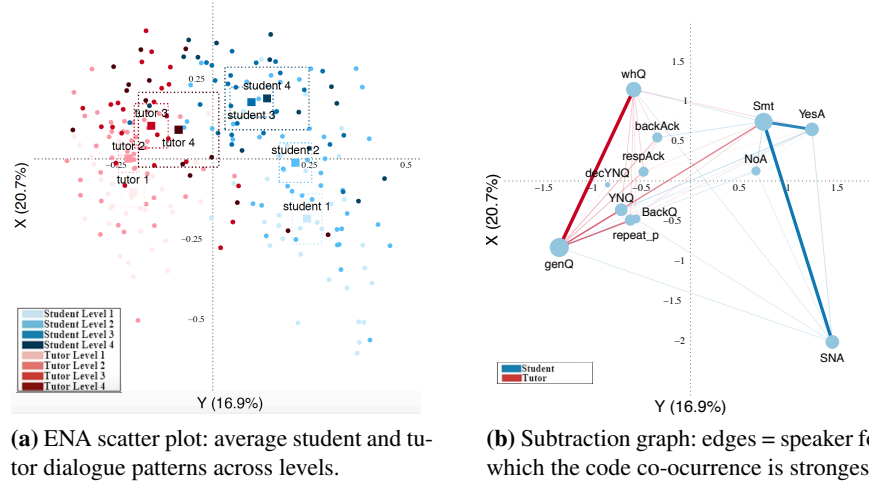
[4] A backchannel is a form of feedback a speaker gives to their interlocutor.

Table 1: DA annotated dialogue examples at Levels 1 (Highest) and 4 (Lowest) in BELC
P = Participant **DA** = Dialogue Act **SPA** = Code-switching in Spanish

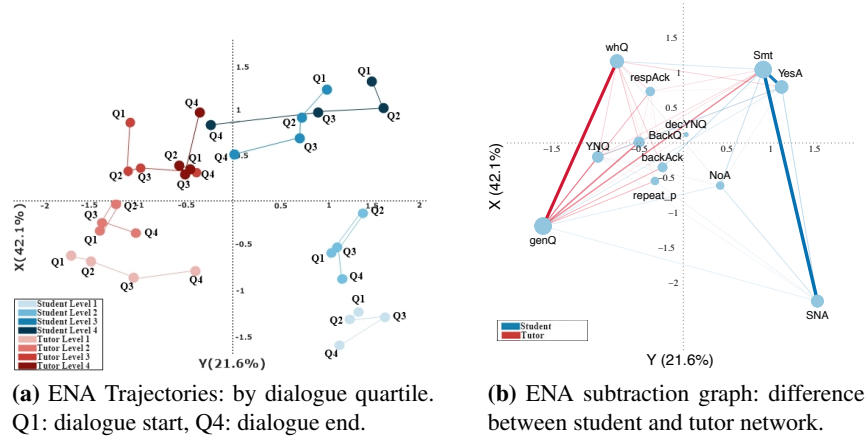| P Level 1	DA	P Level 4	DA	DA key
T do you like the school ?	YNQ	T do you like this school ?	YNQ	**YNQ**: *yes-no-question*
T [- spa] m-entens ?	SPA	S yes .	YesA	**RespAck**: *response*
S 0 [= says nothing] .	SNA	T yes ?	RAck	*acknowledgement*
T "do you like" ?	YNQ	T what are you planning to do next year ?	WhQ	**decYNQ**: *declarative YNQ*
T do you like the school ?	YNQ	S I would like to study zoology .	Smt	**BackQ**: *backchannel-Q*
S xxx .	SNA	T what time did you arrive here this morning	WhQ	**whQ**: *wh-question*
T no si t-agrada l-escola ?	SPA	S this morning ?	GenQ	**GenQ**: *General-Other-Q*
T do you like the school ?	YNQ	T yes .	YesA	**YesA**: *yes answers*
S yes .	YesA	S I ... I am here since eight o'clock .	Smt	**NoA**: *no answers*
T yes ok .	RAck	T uhhuh right quite early .	Smt	**NA**: *non-understanding*
T now what time do you begin in the morning ?	WhQ	T and when will you leave ?	WhQ	**Smt**: *statement*
S 0 [= says nothing] .	SNA	S I ... I finish my time-table in half-past-two .	Smt	**repeat**: *repeat-phrase*
T [- spa] m-entens ?	SPA			**backAck**: *backchannel-*
S [- spa] no .	SPA			*acknowledge*

4 Results & Discussion

To answer RQ1, *full dialogue* is used as stanza, and *speaker* (students and tutors) and *ability level* as unit of analysis. Figure 1(a) shows the projection of individual students' and tutors' networks at different student abilities in DA space. Figure 1(b) shows DA centroids in relation to these axes facilitating interpretation of DA change in Figure 1(a) where it can be seen that speaker means are closer to one another at higher levels of student ability than at lower ones. We also see evidence of tutor movement within DA space (t-tests reveal significant difference between Tutor Level 1 (T1) and Tutor Level 4 (T4): ($D = 1.26$ $p < 0.001$), which we interpret as tutors' adapting their strategy to learner ability. Students show *more* movement in the space across ability level than tutors (t-tests reveal significant difference between S1 and S4: ($D = 1.79$ $p < 0.001$)), indicating that ability influences the sorts of DAs produced, with a more active role (*Wh-Questions (whQ), Response-Acknowledgements (RespAck)* and *Statements (Smt)*) being taken by higher level students. Overall this evidence supports our hypothesis that speakers have similar DA patterns at higher levels of student ability. Figure 1(b) shows that students have more connections between *statements, signal-non-understanding* and *yes-answers* than tutors, who have more connections in general, specifically between *questions, back-channeling* and *repetition*.

(a) ENA scatter plot: average student and tutor dialogue patterns across levels.

(b) Subtraction graph: edges = speaker for which the code co-ocurrence is strongest.

Fig. 1: ENA projection graph: Students and tutors at Higher levels have a position in the space closer to DAs such as *WH-questions*, and *Statements*, whereas at lower levels, they are closer to DAs such as *general questions*, and *Signal-non-understanding (SNA)*.

To answer RQ2, we use *dialogue quartiles* as stanzas and the unit of analysis was again *speaker* and *ability level*. Figure 2(a) shows the trajectory of different groups over

(a) ENA Trajectories: by dialogue quartile. Q1: dialogue start, Q4: dialogue end.

(b) ENA subtraction graph: difference between student and tutor network.

Fig. 2: ENA trajectory graphs. t-tests show significant differences between Q1 & Q4 for each trajectory except Students at Level 1 ($x(D = 0.28\ p = 0.26), y(D = 0.08\ p = 0.74)$). The highest effect sizes were for Tutor Level 1 ($D = 1.49\ p = 0.001$) and Student Level 4 ($D = 1.46\ p = 0.001$).

the four quartiles, each point represents the mean speaker position, in the same DA space as Figure 2(b). Figure 2(a) shows for each level a more similar speaker DA distribution at Q4 than at Q1, supporting our hypothesis of DA convergence over an interaction. At higher levels of student ability, this is more pronounced than for lower levels. Students at higher levels converge to the tutor DA position more than at lower levels (greater distance between Q1 and Q4). We can interpret this as an indication that their greater ability allows them to align more, or that DA usage becomes more diverse with ability. Tutors travel less in the space in general, except with level 1 students, converging to student DA position. We interpret this as evidence of tutor ZPD strategy: converging when the student cannot, and adapting less when they are capable. Sinclair et al. [21] argue higher ability level dialogues become more symmetric, mirroring the symmetric contributions of native speakers. Here, we are able to see that this is the case for interlocutors' use of DAs. While Sinclair et al. [20] found some evidence of alignment at a lexical level, our work shows this at a more abstract level in terms of the conversational dynamics via DAs.

5 Contributions & Conclusions

We contribute a novel method for the analysis of L2 dialogue transcripts, combining the use of Dialogue Act labels with ENA. Our findings support the hypothesis that speakers in L2 dialogue practice exhibit a degree of convergence, both as ability level increases and over the course of a dialogue. The implications of these results are as follows: firstly, a better understanding of tutor adaptation to learners of different ability levels can inform the design of tutoring dialogue systems; secondly, the proposed method can be used to offer formative assessment of learning progression not only to tutors in training, or as a tool for self reflection, but also as a resource for students; and finally, this method can be used by practitioners in learning analytics for the design of new tools across different dialogue modalities. While our proposed method provides evidence of tutor-student convergence, the corpus used is also not large or diverse enough for us to make generalisations about particular dialogue characteristics at certain levels. Accordingly, we limit our interpretation to high-level convergence and adaptation phenomena. An important avenue of future research is to explore the functions that certain DAs perform within dialogue, and the associated difficulty of such acts. The shift in speaker position in DA space suggests a move to using different sorts of DA patterns to better suit the ability of the student. This leads us to hypothesise that certain DA sequences may be more indicative of learner support, while others may capture conversational symmetry.

6 Acknowledgements

Bibliography

[1] Bailey, K.M.: What my efl students taught me. The PAC Journal 1(1), 7–31 (2001)

[2] Birjandi, P., Jazebi, S.: A comparative analysis of teachers' scaffolding practices (2014)

[3] Branigan, H.P., Pickering, M.J., McLean, J.F., Cleland, A.A.: Syntactic alignment and participant role in dialogue. Cognition 104(2), 163–197 (2007)

[4] Costa, A., Pickering, M.J., Sorace, A.: Alignment in second language dialogue. Language and cognitive processes 23(4), 528–556 (2008)

[5] Duplessis, G.D., Clavel, C., Landragin, F.: Automatic measures to characterise verbal alignment in human-agent interaction. In: 18th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL). pp. 71–81 (2017)

[6] Garrod, S., Pickering, M.J.: Alignment in dialogue. The Oxford handbook of psycholinguistics pp. 443–451 (2007)

[7] Hawkes, R.: Learning to Talk and Talking to Learn: How Spontaneous Teacher-learner Interaction in the Secondary Foreign Languages Classroom Provides Greater Opportunities for L2 Learning. Ph.D. thesis, University of Cambridge (2012)

[8] Jurafsky, D., Shriberg, E., Biasca, D.: Switchboard dialog act corpus. International Computer Science Inst. Berkeley CA, Tech. Rep (1997)

[9] Lantolf, J.P.: Second language learning as a mediated process. Language teaching 33(2), 79–96 (2000)

[10] Lantolf, J.P.: Sociocultural theory and second language learning, vol. 78. Oxford University Press (2000)

[11] Mandel, J.: Use of the singular value decomposition in regression analysis. The American Statistician 36(1), 15–24 (1982)

[12] Michel, M.C.: Effects of task complexity and interaction on l2 performance. Second language task complexity: Researching the Cognition Hypothesis of language learning and performance 2, 141–173 (2011)

[13] Muñoz, C.: Age and the rate of foreign language learning, vol. 19. Multilingual Matters (2006)

[14] Pickering, M.J., Garrod, S.: Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences 27(2), 169–190 (2004)

[15] Reitter, D., Keller, F., Moore, J.D.: Computational modelling of structural priming in dialogue. In: Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers. pp. 121–124. NAACL-Short '06, Association for Computational Linguistics, Stroudsburg, PA, USA (2006), http://dl.acm.org/citation.cfm?id=1614049.1614080

[16] Reitter, D., Moore, J.D.: Alignment and task success in spoken dialogue. Journal of Memory and Language 76, 29–46 (2014)

[17] Robinson, P., Gilabert, R.: Task complexity, the cognition hypothesis and second language learning and performance. IRAL-International Review of Applied Linguistics in Language Teaching 45(3), 161–176 (2007)

[18] Samana, W.: Teacher's and students' scaffolding in an efl classroom. Academic Journal of Interdisciplinary Studies 2(8), 338 (2013)

[19] Shaffer, D.W., Hatfield, D., Svarovsky, G.N., Nash, P., Nulty, A., Bagley, E., Frank, K., Rupp, A.A., Mislevy, R.: Epistemic Network Analysis: A Prototype for 21st-Century Assessment of Learning. International Journal of Learning and Media 1(2), 33–53 (2009), doi:10.1162/ijlm.2009.0013

[20] Sinclair, A., Lopez, A., Lucas, C., Gasevic, D.: Does ability affect alignment in second language tutorial dialogue? In: Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue. pp. 41–50 (2018)

[21] Sinclair, A., Oberlander, J., Gasevic, D.: Finding the zone of proximal development: Student-tutor second language dialogue interactions. In: Proc. SEMDIAL 2017 (SaarDial) Workshop on the Semantics and Pragmatics of Dialogue. pp. 107–115 (2017)

[22] Stolcke, A., Ries, K., Coccaro, N., Shriberg, E., Bates, R., Jurafsky, D., Taylor, P., Martin, R., Ess-Dykema, C.V., Meteer, M.: Dialogue act modeling for automatic tagging and recognition of conversational speech. Computational linguistics 26(3), 339–373 (2000)

[23] Van Lier, L.: Interaction in the language classroom: Awareness, autonomy and authenticity. London: Longrnan (1996)

[24] Vygotsky, L.: Zone of proximal development. Mind in society: The development of higher psychological processes 5291, 157 (1987)

[25] Ward, A., Litman, D.: Dialog convergence and learning. Frontiers in Artificial Intelligence and Applications 158, 262 (2007)

[26] Ward, A., Litman, D.: Measuring convergence and priming in tutorial dialog. University of Pittsburgh (2007)