Automatic Measures to Characterise Verbal Alignment in Human-Agent Interaction

G. Dubuisson Duplessis¹, C. Clavel², F. Landragin³
August 15th, 2017

¹ Sorbonne Universités, UPMC Univ Paris 06, CNRS, ISIR, Paris, France
² LTCI, Télécom ParisTech, Université Paris-Saclay, Paris, France
³ Lattice Laboratory, CNRS, ENS, Université de Paris 3, Université Sorbonne Paris Cité, PSL Research University, Paris/Montrouge, France

gdubuisson@telecom-paristech.fr
http://www.dubuissonduplessis.fr
Outline

1. Context: H2020 ARIA VALUSPA Project

2. Background: Convergence and Verbal Alignment

3. Proposition: Automatic Measures to Characterise Verbal Alignment in H-A Interaction

4. Experimentation and Results

5. Conclusion and Perspectives
H2020 European Project: ARIA VALUSPA

Main Features

▶ Artificial Retrieval of Information Assistant
▶ Virtual agent
▶ Multimodal interaction (verbal/non-verbal behaviour)
▶ Adaptation
  - Unexpected situation
  - Socio-emotional state of the user

URL: http://aria-agent.eu/
Outline

1 Context: H2020 ARIA VALUSPA Project

2 Background: Convergence and Verbal Alignment

3 Proposition: Automatic Measures to Characterise Verbal Alignment in H-A Interaction

4 Experimentation and Results

5 Conclusion and Perspectives
Convergence and Verbal Alignment

Communication Accommodation Theory [Gallois et al., 2005]
- Convergence of behaviour occurs both at low-level (e.g., postures, accent and speech rate) and at high-level (e.g., mental, emotional, cognitive)

Interactive Alignment Theory [Pickering and Garrod, 2004]
- Alignment at the lexical, syntactic and semantic levels
- Repetitiveness, routinization and dialogue routines

| Loc. | Utterance |
|------|-----------|
| S₁   | hi i’m sam, nice to meet you what is your name ? |
| H₂   | alex |
| S₃   | how are you doing ? |
| H₄   | i am great |
| S₅   | i really appreciate going fifty fifty with you on clearing out this locker. |

Table 1: Corpus H-A 311 neg1

| Loc. | Utterance |
|------|-----------|
| H₁   | hi |
| S₂   | hi i’m sam, nice to meet you |
| H₃   | nice to meet you i’m erica |
| S₄   | how are you doing ? |
| H₅   | i’m doing good how are you |
| S₆   | pretty good |
| H₇   | good |
| S₈   | i really appreciate going fifty fifty with you on clearing out this locker. |

Table 2: Corpus H-A 376 neg1
Why studying verbal alignment?

**Lessons from H-H interaction**

- Subconscious phenomenon that naturally occurs in H-H dialogues [Pickering and Garrod, 2004]
  - Speakers reuse lexical as well as syntactic structures from previous utterances [Reitter et al., 2006, Ward and Litman, 2007]
- Facilitates successful task-oriented conversations [Nenkova et al., 2008, Friedberg et al., 2012]

**...and what about H-M interaction?**

- Linguistic alignment occurs: users adopt lexical items and syntactic structures used by a system [Brennan and Clark, 1996, Stoyanchev and Stent, 2009, Parent and Eskenazi, 2010, Branigan et al., 2010]
- ...but it is only one-way!
Research Direction

Goal

Provide a virtual agent with the ability to

▶ detect the alignment behaviour of its human interlocutor
▶ align (or not) with the user

Motivation

▶ Natural source of variation in dialogue
▶ Taking into account the socio-emotional behaviour of the user ("social glue")
▶ Adaptation without the need of extensive user profiling

Expected outcomes

▶ Enhancing agent’s believability, likeability and friendliness
▶ Increasing interaction naturalness
▶ Maintaining and fostering user’s engagement [Clavel et al., 2016]
▶ Improving collaboration in task-oriented dialogue
Outline

1 Context: H2020 ARIA VALUSPA Project

2 Background: Convergence and Verbal Alignment

3 Proposition: Automatic Measures to Characterise Verbal Alignment in H-A Interaction

4 Experimentation and Results

5 Conclusion and Perspectives
Proposition

Approach

Providing measures characterising verbal alignment processes based on
- the transcript of dialogue, and
- the shared expressions at the lexical level

![Diagram of proposed framework]

Figure 1: Proposed framework: automatic building of the shared expression lexicon to derive verbal alignment measures.
Automatic Building of the Expr. Lexicon

Shared expression

A surface text pattern at the utterance level that has been produced by both speakers in a dialogue

| A1         | well, that’s an interesting idea. but no, that’s not gonna work for me. |
|------------|-----------------------------------------------------------------------|
| B2         | what will work for you?                                               |
| A3         | what do you think about me getting two chairs and one plate and you getting one chair, one plate, and the clock? |
| B4         | that’s not gonna work for me                                          |

| Expr.                  | Freq. | Init. | ... |
|------------------------|-------|-------|-----|
| that’s not gonna work for me | 2     | A     | ... |
| work for               | 3     | A     | ... |
| me                     | 3     | A     | ... |
| what                   | 2     | B     | ... |
| you                    | 2     | B     | ... |

Figure 2: Main steps to build the dialogue lexicon (inspired from [Dubuisson Duplessis et al., 2017])
### Measures Derived from the Expression Lexicon

| Measure                              | Description                                                                 |
|--------------------------------------|-----------------------------------------------------------------------------|
| Expr. Lexicon Size                   | Number of unique shared expressions in the lexicon (ELS)                    |
| Expr. Variety                        | $EV = \frac{ELS}{\# \text{Tokens}}$                                         |
| Expr. Repetition ($S$)               | $ER_S = \frac{\# \text{Tokens from } S \text{ in an established expr.}}{\# \text{Tokens from } S}$ |
|                                     | $\forall S, ER_S \in [0, 1]$                                                |
| Initiated Expr. ($S$)                | $IE_S = \frac{\# \text{Expr. initiated by } S}{ELS}$                        |
|                                     | $\forall S, IE_S \in [0, 1]$                                                |
Outline

1. Context: H2020 ARIA VALUSPA Project

2. Background: Convergence and Verbal Alignment

3. Proposition: Automatic Measures to Characterise Verbal Alignment in H-A Interaction

4. Experimentation and Results

5. Conclusion and Perspectives
Experimentation Protocol

Protocol

Corpus-based contrastive study to assess the proposed framework and measures

- H-H/A Corpus VS Surrogate Corpora
- H-H Corpus VS H-A Corpus
- Conditions in the H-A Corpus
  - negotiation type (cooperative/competitive)
  - framing (“human operator”/“AI”)
  - gender (male/female agent)
Negotiation Corpora

H-H/A Corpora

- Negotiation task
  - Integrative/win-win, or
  - Distributive/competitive
- 2 settings: H-H, H-A (Woz)
- From [DeVault et al., 2015, Gratch et al., 2016]

The Woz system [DeVault et al., 2015]

- Designed to be as natural as possible
- > 11000 possible utterances

|                | H-H    | H-A (Woz) |
|----------------|--------|-----------|
| Dialogue       | 84     | 154       |
| Utterance      | 10319  | 17125     |
| ... avg (std)  | 122.8 (84.1) | 111.2 (57.5) |
| Token          | 79396  | 90479     |
## Surrogate Corpora

- Break the dynamic of the IAP
- Break the coupling between utterances

| Loc. | Real Utterance | Randomised Utterance |
|------|---------------|----------------------|
| H₁   | hi            | i’m most interested in the chairs |
| S₂   | hi i’m sam , nice to meet you | |
| H₃   | nice to meet you i’m erica | yeah since you won’t budge at all i’d rather do this |
| S₄   | how are you doing ? | |
| H₅   | i’m doing good how are you | why do you want the chairs more than the other items |
|      | [...]         | [...]                |
Results: H-H/A VS Surrogate Corpora

Hypothesis

Dialogue participants should constitute a richer expression lexicon in the H-H/A corpora than what would incidentally happen in the surrogate corpora.

Figure 3: H-H VS surrogate. Expression Variety. Difference is significant ($p < 0.001$).

Figure 4: H-A VS surrogate. Expression Variety. Difference is significant ($p < 0.001$).
Results: H-H/A VS Surrogate Corpora

Hypothesis

Dialogue participants should constitute a richer expression lexicon in the H-H/A corpora than what would incidentally happen in the surrogate corpora.

Results

Observation of richer expression lexicons in the H-H/A corpora than in the surrogate corpora.
**Results: H-H VS H-A Corpora**

**Hypothesis (following [Branigan et al., 2010])**

Verbal alignment differs between H-H and H-A interactions:
- expect more verbal alignment from the human than from the agent (influence by beliefs about the limitations of the agent)

**Figure 3**: Initiated Expressions ($IE_S$). Difference is significant for H-A ($p < 0.001$), not significant for H-H.

**Figure 4**: Expression Repetition ($ER_S$). Difference is significant for H-A ($p < 0.001$), not significant for H-H.
Results: H-H VS H-A Corpora

Hypothesis (following [Branigan et al., 2010])

Verbal alignment differs between H-H and H-A interactions:
▶ expect more verbal alignment from the human than from the agent (influence by beliefs about the limitations of the agent)

Results

Verbal alignment is:
▶ **Symmetrical** in the H-H corpus
▶ **Asymmetrical** in the H-A corpus
  ■ the human participant adopts more Woz-initiated expressions,
  ■ the human participant dedicates more tokens to the repetition of expressions, and
  ■ this asymmetry does not appear when considering the number of tokens produced by each speaker or when considering the proportion of shared vocabulary.
Results: H-A Corpus > Negotiation type

Study

Impact of the negotiation type on verbal alignment indicators
- integrative (win-win)
- distributive (competitive)

Figure 3: Expression Variety (EV). Difference is not significant.

Figure 4: Expression Repetition (ER). Difference is significant ($p < 0.001$).
Results: H-A Corpus > Negotiation type

Study
Impact of the negotiation type on verbal alignment indicators
▶ integrative (win-win)
▶ distributive (competitive)

Results
Competitive negotiation leads to:
▶ longer dialogues,
▶ more verbal alignment (need to verbally align more on (counter-)propositions?)
Outline

1 Context: H2020 ARIA VALUSPA Project

2 Background: Convergence and Verbal Alignment

3 Proposition: Automatic Measures to Characterise Verbal Alignment in H-A Interaction

4 Experimentation and Results

5 Conclusion and Perspectives
Conclusion and Perspectives

▶ Automatic and generic measures of verbal alignment based on sequential pattern mining at the level of surface of text utterances characterising:
  - the routinization process;
  - the degree of repetition between dialogue participants;
  - the orientation of verbal alignment.

▶ Contrasting H-H and H-A verbal alignment (symmetry VS asymmetry)
  - Quantitative confirmation of predictions from previous literature regarding the strength and orientation of verbal alignment in Human-Machine Interaction [Branigan et al., 2010]

▶ Perspectives
  - Online usage in a dialogue system (measures are based on efficient algorithms)
  - Qualitative analysis of verbal alignment differences
  - Confirming results on other comparable H-H/H-A corpora
| Références I |
|--------------|
| **Branigan, H. P., Pickering, M. J., Pearson, J., and McLean, J. F. (2010).**  
*Linguistic alignment between people and computers.*  
*Journal of Pragmatics, 42(9):2355–2368.* |
| **Brennan, S. E. and Clark, H. H. (1996).**  
*Conceptual pacts and lexical choice in conversation.*  
*Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(6):1482.* |
| **Clavel, C., Cafaro, A., Campano, S., and Pelachaud, C. (2016).**  
*Fostering user engagement in face-to-face human-agent interactions: a survey.*  
*In Toward Robotic Socially Believable Behaving Systems-Volume II, pages 93–120. Springer.* |
| **DeVault, D., Mell, J., and Gratch, J. (2015).**  
*Toward natural turn-taking in a virtual human negotiation agent.*  
*In AAAI Spring Symposium on Turn-taking and Coordination in Human-Machine Interaction.*  
*AAAI Press, Stanford, CA.* |
| **Dubuisson Duplessis, G., Charras, F., Letard, V., Ligozat, A.-L., and Rosset, S. (2017).**  
*Utterance Retrieval based on Recurrent Surface Text Patterns.*  
*In 39th European Conference on Information Retrieval (ECIR), pages 199–211, Aberdeen, United Kingdom.* |
| **Friedberg, H., Litman, D., and Paletz, S. B. (2012).**  
*Lexical entrainment and success in student engineering groups.*  
*In Spoken Language Technology Workshop (SLT), pages 404–409. IEEE.* |
Gallois, C., Ogay, T., and Giles, Howard, H. (2005). Communication accommodation theory: A look back and a look ahead. W. Gudykunst (red.): Theorizing about intercultural communication. Thousand Oaks, CA: Sage, pages 121–148.

Gratch, J., DeVault, D., and Lucas, G. (2016). The benefits of virtual humans for teaching negotiation. In International Conference on Intelligent Virtual Agents (IVA), pages 283–294. Springer.

Nenkova, A., Gravano, A., and Hirschberg, J. (2008). High frequency word entrainment in spoken dialogue. In Proceedings of the 46th annual meeting of the association for computational linguistics on human language technologies (ACL-HLT): Short papers, pages 169–172. Association for Computational Linguistics.

Parent, G. and Eskenazi, M. (2010). Lexical entrainment of real users in the let’s go spoken dialog system. In INTERSPEECH, pages 3018–3021.

Pickering, M. J. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. Behavioral and brain sciences, 27(02):169–190.
Références III

Reitter, D., Keller, F., and Moore, J. D. (2006). Computational modelling of structural priming in dialogue. In Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL (NAACL-HLT): Short Papers, pages 121–124. Association for Computational Linguistics.

Stoyanchev, S. and Stent, A. (2009). Lexical and syntactic priming and their impact in deployed spoken dialog systems. In Proceedings of the Human Language Technology Conference of the North American Chapter of the ACL (NAACL-HLT): Short Papers, pages 189–192. Association for Computational Linguistics.

Ward, A. and Litman, D. J. (2007). Automatically measuring lexical and acoustic/prosodic convergence in tutorial dialog corpora. In Speech and Language Technology in Education (SLaTE2007), pages 57–60.