Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Mutual Gaze and Linguistic Repetition in a Multimodal Corpus

Anaïs Murat, Maria Koutsombogera, Carl Vogel

LREC 2022
09/05/2022

Agenda

- **Introduction**
  - Alignment & Linguistic Repetitions
  - Gaze & Mutual Gaze
  - What link between both?
- **Methods**
- **Results & Discussion**
- **Conclusion**

INTRODUCTION

Introduction

Alignment & Linguistic Repetition

- *How to know whether people understand each other?*
- **Interactive Alignment Model: shared situational model** (Pickering & Garrod, 2004)
  - Development of routines to disambiguate terms in context.
- **No universal method to study alignment (Doyle and Frank, 2016):**
  - *Which levels of representation? Punctuation in transcripts or not?*
    - **Syntactic alignment** more relevant than lexical alignment(= topic) (Reitter & Moore (2007).
    - Study of **5 levels of representation** (Reverdy et al., 2020):
    - In-isolation levels: token, lemma, and part-of-speech (POS)
    - Paired levels: token + POS, and lemma + POS

Introduction

Gaze & Mutual Gaze (MG)

- **Theory of mind** and **gaze (Emery, 2000)**
- **Shows where the attention of the person is**
- **Mutual gaze can:**
  - Help manage turn-taking
  - Initiate social interaction (Cary, 1978, in Pfeiffer et al. (2013))
  - Show willingness to pursue the conversation (Jokinen et al., 2010)
- **As opposed to averted gaze:**
  - reduces cognitive load (Jording et al., 2018)
  - willingness not to continue the interaction in the same terms (Jokinen et al., 2010).

Introduction

What possible link between both?

- **Mutual Gaze:** cue of the theory of mind, focus, and manage interactions
- **Linguistic repetitions:** show alignment which demonstrates understanding

- Both seem to inform the interaction and help with its progression.
- Our starting point hypotheses are that **mutual gaze is in greater evidence** at **times of mutual understanding** than times **without mutual understanding.**

Introduction

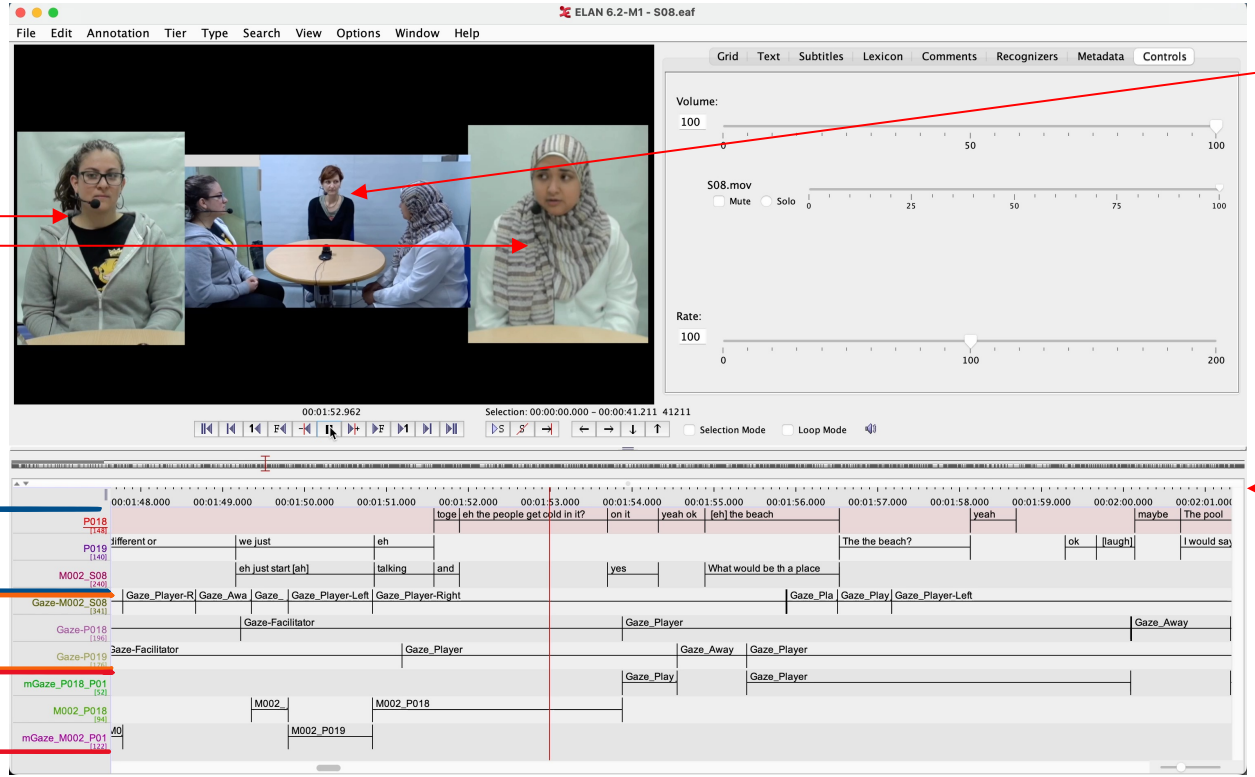
Goals

- **Explore ways to inform conversation and Natural Language Understanding in interactions**
- **Investigate a possible relation between mutual gaze and linguistic repetitions**
- **A contribution to the method to measure alignment in real-time**

METHODS

Methods

Data Collection: the Multisimo Corpus (Koutsombogera & Vogel, 2018)



Methods

Data Collection: the Multisimo Corpus

HOW TO ALIGN THESE DATA IN A 2D ARRAY?

Methods

Alignment of the data

1) Between **turns** and **mutual gazes**

2) Adding levels of representation (tokens (T), lemmas (L) , and parts-of-speech (POS))
   - TreeTagger (Schmid, 1994)

3) Adding counts of repetitions: other-repetitions & self-repetitions

Turn	Mutual Gaze
Hello	MG1
Hello, how are you?	MG1
Hello, how are you?	MG2
Good.	NONE

Methods

Alignment of the data

1) Between **turns** and **mutual gazes**
2) Adding **levels of representation** (tokens (T), lemmas (L) , and parts-of-speech (POS)) - TreeTagger (Schmid, 1994)
3) Adding counts of repetitions: other-repetitions & self-repetitions

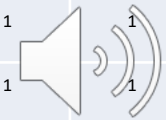
Turn	Mutual Gaze	Speech tags		
		Token	Lemma	POS
Hello	MG1	hello	hello	UH
Hello, how are you?	MG1	hello; how; are; you; ?	hello; how; be; you; ?	UH; RB; VBP; PP; ?
Hello, how are you?	MG2	hello; how; are; you; ?	hello; how; be; you; ?	UH; RB; VBP; PP; ?
Good.	NONE	good; .	good; .	JJ; .

Methods

Alignment of the data

1) Between **turns** and **mutual gazes**
2) Adding **levels of representation** (tokens (T), lemmas (L) , and parts-of-speech (POS)) - TreeTagger (Schmid, 1994)
3) Adding **counts of repetitions: other-repetitions** & self-repetitions

Turn	Mutual Gaze	Speech tags			Counts of repetition (uni-grams)									
					Other-Repetitions									
					Punctuation					No Punctuation				
		T	L	POS	T	L	POS	T+POS	L+ POS	T	L	POS	T+POS	L+ POS
hello	MG1	hello	hello	UH	0	0	0	0	0	0	0	0	0	0
Hell..	MG1	how..	how..	RB..	1	1	1	1	1	1	1	1	1	1
Hell..	MG2	how..	how..	RB..	1	1	1	1	1	1	1	1	1	1

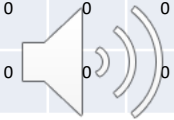


Methods

Alignment of the data

1) Between **turns** and **mutual gazes**
2) Adding **levels of representation** (tokens (T), lemmas (L) , and parts-of-speech (POS)) - TreeTagger (Schmid, 1994)
3) Adding **counts of repetitions:** other-repetitions & **self-repetitions**

…	Speech tags			Counts of repetition																	
	T	L	POS	Other-Repetitions										Self-Repetitions							
				?!.					ø					?!.					ø					
				T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS
…	…	…	…	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
…	…	…	…	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0
…	…	…	…	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0



Methods

Alignment of the data

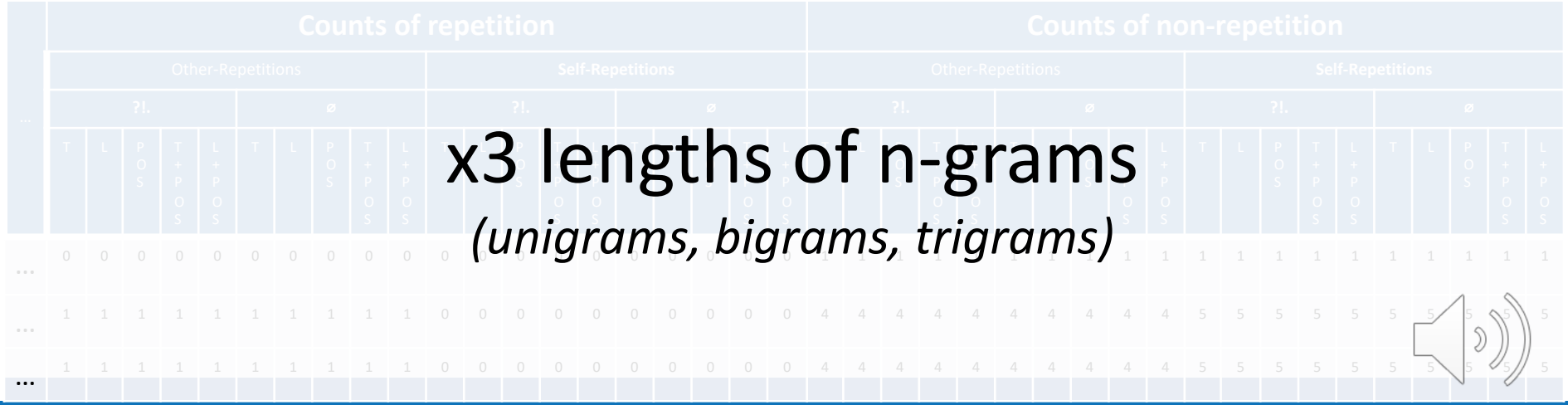
1) Between **turns** and **mutual gazes**
2) Adding **levels of representation** (tokens (T), lemmas (L) , and parts-of-speech (POS)) - TreeTagger (Schmid, 1994)
3) Adding **counts of repetitions:** other-repetitions & **self-repetitions…** and non-repetitions.

…	Counts of repetition																				Counts of non-repetition																	
	Other-Repetitions										**Self-Repetitions**										Other-Repetitions										**Self-Repetitions**									
	?!.					**ø**					**?!.**					**ø**					**?!.**					**ø**					**?!.**					**ø**				
	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS
…	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
…	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5
…	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	4	4	4	4	4	4	4	4	4	4	5	5	5	5	5	5	5	5	5	5

Methods

Alignment of the data

1) Between **turns** and **mutual gazes**
2) Adding **levels of representation** (tokens (T), lemmas (L) , and parts-of-speech (POS))
   - TreeTagger (Schmid, 1994)
3) Adding **counts of repetitions:** other-repetitions & **self-repetitions…** and non-repetitions.



Methods

Alignment

1) Between **turns** and **mutual gazes**
2) Adding **levels of representation** (tokens (T), lemmas (L) , and parts-of-speech (POS))
3) Adding **counts of repetitions** & **self-repetitions**

Is there a relation between **linguistic repetitions** and the $\begin{cases} \boldsymbol{presence} \\ \boldsymbol{duration} \end{cases}$ **of mutual gaze**?

Turn	Mutual Gaze	Speech tags			Counts of repetition																
		T	L	POS	Other-Repetitions										Self-Repetitions						
					?I.					ø					?I.					ø		
					T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS	T	L	POS	T+POS	L+POS

RESULTS & DISCUSSION

Results & Discussion

The Presence of Mutual Gaze & Linguistic Repetition

Punctuation

Other-Repetitions

Level	n-grams,$n=1$ χ^2	p	n-grams,$n=2$ χ^2	p	n-grams,$n=3$ χ^2	p
Token	36.272	1.72E-09	5.4979	0.01904	2.4851	0.1149
POS	249.9	<2.2e-16	59.463	1.25E-14	12.664	0.0003728
Lemma	46.374	9.77E-12	5.2342	2.22E-02	1.1553	0.2824
Token+POS	12.801	0.0003464	0.028789	0.8653	0.20761	0.6486
Lemma+POS	10.375	0.001277	0.0041934	0.9484	0.17917	0.6721

Self-Repetitions

Level	n-grams,$n=1$ χ^2	p	n-grams,$n=2$ χ^2	p	n-grams,$n=3$ χ^2	p
Token	70.724	<2.2e-16	6.595	0.01023	2.99E-28	1
POS	317.54	<2.2e-16	80.628	<2.2e-16	14.335	0.000153
Lemma	97.116	<2.2e-16	12.707	0.0003643	0.18828	0.6643
Token+POS	1.4719	0.2251	0.18024	0.6712	1.1845	0.2764
Lemma+POS	0.3333	0.3333	0.92186	0.337	0.18594	0.6663

No Punctuation

Other-Repetitions

Level	n-grams,$n=1$ χ^2	p	n-grams,$n=2$ χ^2	p	n-grams,$n=3$ χ^2	p
Token	16.138	5.89E-05	8.809	0.002996	2.11	0.1463
POS	243.87	<2.2e-16	61.479	4.48E-15	13.248	0.0002728
Lemma	25.58	4.23E-07	7.5885	0.005874	1.1872	0.2759
Token+POS	13.092	0.0002965	0.11381	0.7358	0.030156	0.8621
Lemma+POS	11.213	0.0008121	0.50578	0.477	0.065683	0.7977

Self-Repetitions

Level	n-grams,$n=1$ χ^2	p	n-grams,$n=2$ χ^2	p	n-grams,$n=3$ χ^2	p
Token	53.419	2.69E-13	4.7997	0.02847	0.24214	0.6227
POS	306.65	<2.2e-16	83.497	<2.2e-16	17.761	2.50E-05
Lemma	79.783	<2.2e-16	10.92	0.0009475	1.3245	0.2498
Token+POS	6.3182	0.01195	0.97688	0.323	1.0491	0.3057
Lemma+POS	5.0635	0.02444	0.23492	0.6279	8.45E-27	1

Results & Discussion

The Presence of Mutual Gaze & Linguistic Repetition

	?!.	∅
Other-Repetitions		
Self-Repetitions		

Is there a relation between **linguistic repetitions** and the {*presence* ✔ / *duration*} **of mutual gaze**?

Results & Discussion

The Presence of Mutual Gaze & Linguistic Repetition

	?!.	∅
Other-Repetitions		
Self-Repetitions		

Is there a relation between **linguistic repetitions** and the {~~***duration***~~ / ***presence*** ✔} **of mutual gaze**?

Other-Repetitions	n-grams, $n=1$		n-grams, $n=2$		n-grams, $n=3$	
Level	χ^2	p	χ^2	p	χ^2	p
Token	36.272	1.72E-09	5.4979	0.01904	2.1951	0.1149
POS	249.9	<2.2e-16	59.463	1.23E-14	13.248	0.0002728
Lemma	46.374	9.77E-12	5.2	0.02	1.1553	0.2826
Token+POS	12.801	0.0003464	0.028789	0.8652	0.11381	0.7358
Lemma+POS	10.375	0.001277	0.0041934	0.9484	0.17917	0.6721

Other-Repetitions	n-grams, $n=1$		n-grams, $n=2$		n-grams, $n=3$	
Level	χ^2	p	χ^2	p	χ^2	p
Token	16.138	5.89E-05	8.809	0.002996	2.11	0.1463
POS	243.87	<2.2e-16	61.479	4.48E-15	13.248	0.0002728
Lemma	29.6	5.23E-07	7.5885	0.005874	1.1872	0.2759
Token+POS	13.092	0.0002965	0.11381	0.7358	0.030156	0.8621
Lemma+POS	11.213	0.0008121	0.50578	0.477	0.065683	0.7977

Self-Repetitions	n-grams, $n=1$		n-grams, $n=2$		n-grams, $n=3$	
Level	χ^2	p	χ^2	p	χ^2	p
Token	70.724	<2.2e-16	6.595	0.01023	2.99E-28	1
POS	317.54	<2.2e-16	80.628	<2.2e-16	14.335	0.000153
Lemma	97.116	<2.2e-16	12.707	0.0003643	0.18828	0.6643
Token+POS	1.4719	0.2251	0.18024	0.6712	1.1845	0.2764
Lemma+POS	0.3333	0.3333	0.92186	0.337	0.18594	0.6663

Self-Repetitions	n-grams, $n=1$		n-grams, $n=2$		n-grams, $n=3$	
Level	χ^2	p	χ^2	p	χ^2	p
Token	53.419	2.69E-13	4.7997	0.02847	0.24214	0.6227
POS	306.65	<2.2e-16	83.497	<2.2e-16	17.761	2.50E-05
Lemma	79.783	<2.2e-16	10.92	0.0009475	1.3245	0.2498
Token+POS	6.3182	0.01195	0.97688	0.323	1.0491	0.3057
Lemma+POS	5.0635	0.02444	0.23492	0.6279	8.45E-27	1

CONCLUSION

Conclusion

Main Points

- **Mutual gaze presence and linguistic repetition:**
  - All lengths of n-grams were shown significant.
  - All level of representation were shown significant
  - Positive correlation for token, lemma and POS in isolation.
  - Negative correlation for paired levels of representations (Token + POS, Lemma + POS).
- **Improvement of the method:**
  - Evaluation of the tagger on speech data
  - Real-time alignment
  - Punctuation led to less significancy

Conclusion

Weaknesses

- Mainly significant for uni- and bigrams: a question of size?
- No significant results for duration
- Individuation per turn and mutual gaze:
  - Excessive weight for turns containing mutual gazes.
  - What about the other types of gaze?
- Mutual Gaze methods: who are truly involved in them?

Acknowledgement

This work was conducted with the financial support of the **Science Foundation Ireland Centre for Research Training in Artificial Intelligence under Grant No. 18/CRT/6223**, and was supported by **the GEHM research network (Independent Research Fund Denmark grant 9055-00004B).**

Bibliography

- Brugman, H. and Russel, A. (2004). Annotating multimedia/multi-modal resources with ELAN. In *Proceedings of the Fourth International Conference on Language Resources and Evaluation (LREC 2004)*, pages 2065–2068, Lisbon, Portugal. European Language Resources Association (ELRA).
- Cary, M. S. (1978). The role of gaze in the initiation of conversation. *Social Psychology*, 41(3):269–271.
- Doyle, G. and Frank, M. C. (2016). Investigating the Sources of Linguistic Alignment in Conversation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 526–536, Berlin, Germany. Association for Computational Linguistics
- Emery, N. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6):581–604, August

Bibliography

- Jokinen, K., Harada, K., Nishida, M., and Yamamoto, S. (2010). Turn-Alignment Using Eye-Gaze and Speech in Conversational Interaction. *Interspeech*, pages 2018–2021.
- Jording, M., Hartz, A., Bente, G., Schulte-Ruther, M., ¨ and Vogeley, K. (2018). The “Social Gaze Space”: A Taxonomy for Gaze-Based Communication in Triadic Interactions. *Frontiers in Psychology*, 9:226, February
- Koutsombogera, M. and Vogel, C. (2018). Modeling collaborative multimodal behavior in group dialogues: The multisimo corpus. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, pages 2945– 2951.
- Pickering, M. J. and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02):169–190, April.
- Reitter, D. and Moore, J. D. (2007). Predicting success in dialogue. In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 808–815. Association for Computational Linguistics

Bibliography

- Reverdy, J., Koutsombogera, M., and Vogel, C. (2020). Linguistic Repetition in Three-Party Conversations. In Anna Esposito, et al., editors, *Neural Approaches to Dynamics of Signal Exchanges*, volume 151, pages 359–370. Springer Singapore, Singapore. Smart Innovation, Systems and Technologies.
- Schmid, D. (1994). Probabilistic part-of-speech tagging using decision trees. *Proceedings of Interna- tional Conference on New Methods in Language Processing*, pages 44–49

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Thank You