DIFFERENCES IN THE COMMUNICATIVE INTENT AMONG SPEAKERS IN FALSE AND (PRESUMED) TRUE CONFESSIONS

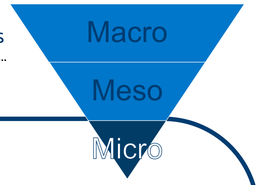
Guadalupe Blanco-Velasco & Kimberley A. Clow

Ontario
Social Sciences and Humanities Research Council of Canada
Conseil de recherches en sciences humaines du Canada

INTRODUCTION

Interview approaches and types of false confessions

Techniques and strategies



MICRO-LEVEL: Questions/Answers and Content

PREVIOUS RESEARCH ON INVESTIGATIVE INTERVIEWING

- Focused on one level (i.e., type of questions; e.g., Lamb et al., 2007) without integrating across levels.
- Current micro-level analyses do not capture the full continent (i.e., communicative intention) and content (i.e., the topic) of entire speaking turns (e.g., Waterhouse et al., 2019).
- Rely on subjective interpretations (e.g., "appropriateness"; Griffiths & Milne, 2006).

PREVIOUS RESEARCH ON FALSE CONFESSIONS (vs TRUE CONFESSIONS)

- Difficulties distinguishing between types of confessions due to similar quality of content (e.g., Appleby et al., 2013).
- Focused on specific elements in statements but a lack of research with interview transcripts (i.e., pronouns and conjunctions; Rizelli et al., 2021).

AIM

Analyse the first element of the Micro-level analysis—communicative intention (i.e., continent) — and identify potential differences between false and (presumed) true confessions.

METHOD

SAMPLE

20 complete false confessions.
20 complete (presumed) true confessions.

SOURCES

National Registry of Exonerations.
Publicly available on the internet.
Researchers.

DATA PREPARATION

Whisper AI transcription (e.g., YouTube videos).
Python 3 scripts (e.g., PDF to Excel).
Manual transcription, review, and anonymization.

LARGE LANGUAGE MODEL TRAINING

5 false and 5 (presumed) true confessions:
Train set: 4,597 speaking turns and 7,046 units.
Validation set: 575 speaking turns and 874 units.
Test set: 675 speaking turns and 1,146 units.

SEGMENTATION MODEL

Divides speaking turns into units (Auld & White, 1956).
Base model: LED (Beltagy et al., 2020).
Validation F1: 95.7%.
Test accuracy: 89.1% general and 85% exact matches (92.5% and 88.4% respectively, with post-processing rules).

CONTINENT CLASSIFICATION MODEL

Base model: RoBERTa (Liu et al., 2019).
Validation F1: 94%.
Test accuracy: 93% general and 95% exact matches.

This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET:www.sharcnet.ca) and Digital Research Alliance of Canada (https://alliancecan.ca/en)

CONTINENT CODEBOOK

Invitation: Questions to obtain a free recall (e.g., "What happened next?").

Probing: Questions to elicit elaboration, clarification, or detail about a specific topic (e.g, "How was the car?").

Yes/No: Questions to get a yes/no answer (e.g., "Were you at home?").

Option-posing: Questions to get an answer based on at least two options (e.g., "Was it blue or red?").

Check questions (Othman, 2010)**:** Questions to confirm knowledge, progress the conversation, or get the listener's attention (e.g., "I want you to be truthful, okay?").

Backchannel (Schegloff, 1982; Yngve, 1970)**:** Aimed to indicate active listening or acknowledgment without taking a full turn (e.g., "okay", "mmh-mm").

Elliptical (Gunter, 1963)**:** Units at the beginning of the speaking turn that lack explicit elements (subject/verb) (e.g., "The house").

Assertive: Commits the speaker to the truth of the expressed proposition (e.g., "I was there").

Directive: The speaker attempts to get the listener to do something (e.g., "Just tell me then").

Commissive: Commits to a future action that can be carried out without the involvement of the listener (e.g., "I'll get into that in a second").

Expressive: The speaker expresses psychological or emotional states or reactions (e.g., "I don't mean to interrupt").

Yes/No answer: Units at the beginning of the speaking turns with "yes" or "no".

Incomplete: Units that are syntactically or semantically incomplete and do not allow for a complete independent clause (e.g., "I was--").

Unintelligible: Units with unintelligible content that do not allow for a complete independent clause (e.g., "He went (unintelligible) when it was early").

(Griffiths & Milne, 2006; Korkman et al., 2006; Lamb et al., 2007; Searle, 1976; Waterhouse et al., 2016)

RESULTS

COMMUNICATIVE INTENT BY CONFESSION TYPE

Continent	FALSE (n/%)	TRUE (n/%)	|StdRes|	Sig.
Yes/No	3,847 (15.2 %)	3,131 (11.1 %)	|14.12|	< .001 *
Probing	2,345 (9.23%)	1,751 (6.2 %)	|13.32|	< .001 *
Yes/No answer	1,627 (6.4 %)	1,455 (5.2 %)	|6.33|	< .001 *
Check questions	459 (1.8 %)	436 (1.5 %)	|2.43|	.015 *
Commissive	434 (1.7%)	328 (1.1 %)	|5.40|	< .001 *
Assertive	10,778 (42.8 %)	13,182 (46.9 %)	|9.54|	< .001 *
Backchannel	811 (3.2 %)	1 834 (6.5 %)	|17.55|	< .001 *
Incomplete	740 (2.9 %)	1,073 (3.8 %)	|5.60|	< .001 *
Expressive	485 (1.9 %)	755 (2.7 %)	|5.82|	< .001 *
Unintelligible	404 (1.6 %)	646 (2.3 %)	|5.76|	< .001 *
Invitation	136 (0.5 %)	132 (0.5 %)	|1.14|	0.253
Directive	742 (2.9 %)	796 (2.8 %)	|0.78|	0.435
Elliptical	2,067 (8.2 %)	2,223 (7.9 %)	|1.25|	0.21
Option-posing	321 (1.3 %)	369 (1.3 %)	|0.39|	0.694

Note. |StdRes| > 1.96 ($p < .05$); Overall association: $x^2(13, N = 53,307) = 851.25$, $p < .001$; Cramer's V = .12, 95% CI [.11, .13].

DISCUSSION

Significant differences between types of confession

False confessions

- Characterized by closed questions and commissive units.

(Presumed) true confessions

- Included more listener signals (backchannel), assertive statements, expressive units, and fragments.

Invitations were rare in each type of interview.

This study shows **significant differences at the micro-level** without considering abstract elements (e.g., suggestiveness).

Further research should aim to develop a tool (i.e., LLM pipeline) that can **estimate the probability** that a confession is false or (presumed) true.

References

Contact