Simon Fraser University
Language and Brain Lab

Multi-speaker experimental designs: Methodological Considerations

Offrede et al., 2021

KEYWORDS: MULTIMODALITY, RESEARCH METHODS, SPEAKER PREPARATION, DATA PROCESSING

Research Aim

- "To contribute to the advancement in studies of **multimodal** language use in **naturalistic settings**"
    - i.e., find ways to study the different facets of language (speech, auditory cues, visual gestures, etc.) in everyday settings
- Summarizes guidelines on experimental designs involving multimodality, especially in multi-speaker experiments

Methodology

- 29 researchers w/ previous experience of dual-speaker experimental design filled out **questionnaire**
- 13 questions in questionnaire, containing the following topics:
    - Experimental tasks & techniques used
    - Greatest challenges of dual-speaker experiments
    - Resolving problems pertaining to participant logistics
    - Data collection & processing
    - Useful recommendations for researchers
- 3 other experts **interviewed** for further information on data collection

Speaker Recruitment

- Various resources to recruit participants for dual-speaker studies:
    - University's participant database
    - Mailing lists
    - Poster advertisements on campus
    - Social media
- Factors to potentially consider when pairing participants:
    - Gender
    - Personality traits
    - Familiarity (friends or strangers)
    - Bias in representation
        - **WEIRD** (Western, Educated, Industrialized, Rich, Democratic)

Task-Oriented Experiments - I

1. **Tangram Task**
   - 2 speakers given identical set of images, but arranged in different orders
   - Cannot directly look at partner's set of images
   - Rearrange images in matching order by communicating w/ partner
   - Many variations based on this experimental setup
2. **4 Images, 1 Arrow**
   - Based on Tangram Task; speakers split into "Director" & "Matcher" role
   - **Director**: Given 4 images w/ 1 arrow pointing at one of the images
   - **Matcher**: Given 1 of the 4 images
   - Determine whether Matcher's image matches picture Director's arrow is pointing to

Task-Oriented Experiments - II

3. **Diapix**
   - Based on Tangram Task; speakers given nearly identical images, but minor details differ
   - Cannot directly look at partner's set of images
   - Identify all differences between both images by communicating w/ partner
   - Perceived as **most balanced setup** for equal participation between both speakers

(a)

(b)

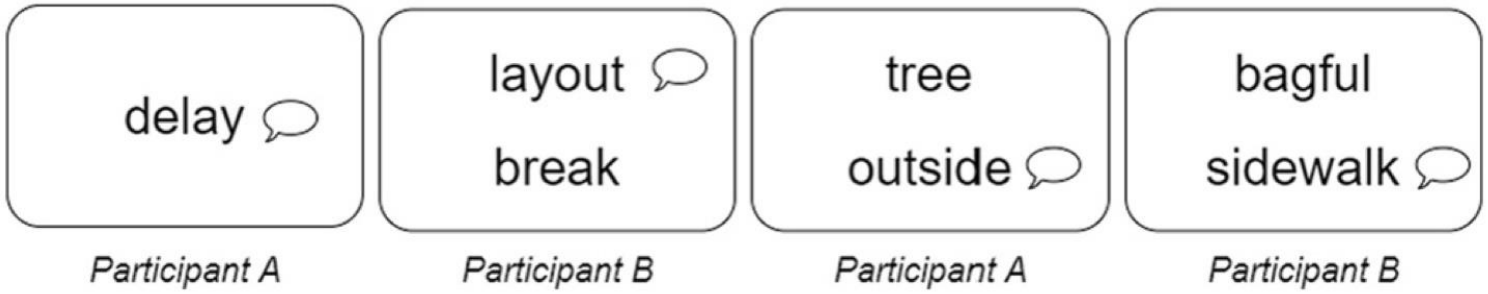
(Baker & Hazan, 2011)

Task-Oriented Experiments - III

4. **Map Task**
   - Based on Tangram Task; speakers given nearly identical maps
   - **Giver**: Map contains dotted path
   - **Follower**: Map does not contain dotted path
   - Cannot directly look at partner's map
   - Follower must trace path on map w/ pencil based on Giver's instructions
   - Speakers are not informed that maps are slightly different for extra challenge
   - Variation of Map Task: one speaker views image & gives verbal instructions to partner on how to draw picture

Task-Oriented Experiments - IV

5. **Domino Task**
   - Controlled interactive task
   - Speakers shown pairs of words on computer screen
   - Within pair, must say word w/ first syllable matching last syllable of word previously said by partner
   - Can be used to study **phonetic convergence**: how speech changes depending on person being spoken to



Free Conversation Experiments

- Can be structured or unstructured
- Some experiments suggest conversation topics, which may help guide untalkative participants
    - e.g., hobbies, food, holidays
- Structured conversations help control extent of emotional reactions & gestures for sensitive topics
    - e.g., politics, religion, life/death

Techniques: Audio

- Audio & video most common experimental setup for 2 or more speakers
- Often either audio-only, or both audio & video recordings
- **Stereo** recordings preferred over mono recordings
    - Allows input from 2 microphones simultaneously instead of 1
- **Cardioid** microphones recommended: captures input from front of mic better than back (prevents mic from gathering partner's speech if in close proximity)
- **Head-mounted** mics may be more accurate than stand-mounted mics due to consistent distance between mic & mouth

Techniques: Video

- Ensure all speakers are in frame w/ gestures & facial expressions visible
- Video not recommended for faster movements (e.g., lips)
    - Consider motion capture as an alternative to video
- Can analyze body movements from videos for conversation gestures using pixel changes or deep learning
    - **Pixel Changes**: Calculates movement speed/direction from difference in brightness of pixels in each frame
    - **Deep Learning**: Determines location of body parts in frame & can identify speaker that each body part belongs to
        - More suitable for multi-speaker studies because it eliminates need to worry about overlapping images

Techniques: Electroencephalography (EEG)

- Extremely common neuroimaging technique
- Electrodes placed along scalp & records surface-level brain activity
- 2 approaches to EEG usage in dual-speaker experiments:
    - Separately studying EEG signals of each speaker
    - **Neural coupling**: neurons fired similarly in brain of listener when speaker tells a story
- Computational analyses of EEG data possible through Python & MATLAB programming languages

Other Techniques

Respiratory Inductance Plethysmography (RIP)

- Becoming more common in multi-speaker studies
- 2 respiratory belts wrapped around speaker's upper body (ribcage/abdomen)
- Used to measure change in lung volume when speaking

Electromagnetic Articulography (EMA)

- Solves issue of overlapping audio recordings in multi-speaker environment by recording articulatory movements
- Uses electromagnetic induction – attach sensors w/ coil to mouth (e.g., tongue, lips, jaw)

Most Relevant Information

- **Task-oriented** & **free conversation** most popular methods for speech elicitation in multi-speaker studies
    - **Tasks:** Domino, Diapix, variations of Tangram Task
    - **Free Conversation:** Structured w/ topics or unstructured

- Equipment considerations for audio & video techniques
    - Stereo recordings, cardioid/head-mounted microphones

Summary of techniques

TABLE A4 Summary of information concerning acoustic, video and respiration recordings

	Acoustics	Video	Respiration
Time to prepare equipment	5–10 min	5–10 min	5 min
Time to prepare participants	5–10 min	5–10 min	10 min
Lab members to prepare participants	1–2	1–2	1
Lab members to record data	1	1	1–2
Computers to record data	1	1	1–2
Postprocessing software	Audacity; SoX; MATLAB; Python; Praat	FFmpeg; R; MATLAB	Python; R
Annotation software	Praat; ELAN; MATLAB; Excel	ELAN; FFmpeg; Praat; Excel	Praat; ELAN

TABLE A5 Summary of information concerning eye-tracking, motion capture and EMA recordings

	Eye tracking	Motion capture	EMA
Time to prepare equipment	5–10 min	10 min	1 hr
Time to prepare participants	5–10 min	5 to 10 min	30–45 min per participant
Lab members to prepare participants	1–2	1–2 per participant	2 per participant
Lab members to record data	1	1	2–3
Computers to record data	2	1	2–3
Postprocessing software	R; MATLAB	Python; MATLAB	Python; MATLAB
Annotation software	ELAN; MATLAB; Excel	Praat; ELAN	Python; MATLAB (e.g. MView); VisArtico

TABLE A6 Summary of information concerning EEG, skin conductance and fNIRS recordings

	EEG	Skin conductance	fNIRS
Time to prepare equipment	20 min	5–10 min	30 min
Time to prepare participants	ca. 30 min per participant	5–10 min	20 min
Lab members to prepare participants	2 per participant	1–2	2 per participant
Lab members to record data	1–2	1–2	2
Computers to record data	2–3	2	6
Combine data	Python; MATLAB	Python; MATLAB	MATLAB
Postprocessing software	Python; MATLAB	Python; MATLAB	MATLAB
Annotation software	MATLAB	Python; MATLAB; ACKNOWledge	MATLAB; Equipment's software