Data analysis plan – Patricc 1.0

1. Preprocessing
   1. Trim data – From start cue to end cue.
2. Create new features
   1. Conversational turns
      1. 4 types – PC (parent child), CP (child parent), PCP (Parent child parent), CPC (child parent child).
      2. Counted if there is a gap smaller the 5 s between utterances of child and parent.

לקחתי אם יש הפרש של שניה בהתחלה והתחיל לפחות 5 שניות אחרי שהראשון סיים, אם היה מספר אמירות לאחר האינטרקציה נלקחה הראשונה בלבד.

* + 1. 3 utterance turn are counted as 2 turns. PC and CP are counted as one conversational turn. CPC and PCP are counted as 2 conversational turns.

חושב כמו סעיף קודם רק במקום שזה ילד והורה זה אינטרקציה הורה ילד ו ינטקציה ילד הורד

* 1. Gaze
     1. Joint attention- times that the parent and child are both looking at the same thing.

נספר כאשר תחילת ההסתכלות של אדם 1 היא בתוך החלון הסתכלות של אדם 2, זמן סיום האינטרקציה הוא כאשר אחד מהם מפסיק להסתכל

* + 1. Joint attention at robot – times that the parent and child are both looking at the robot.
    2. Joint attention at props- times that the parent and child are both looking at the props.
    3. Joint attention at tablet- times that the parent and child are both looking at the tablet.
    4. Mutual gaze- times that the parent and child are looking at each other.

נספר כאשר תחילת ההסתכלות של אדם 1 היא בתוך החלון הסתכלות של אדם 2, זמן סיום האינטרקציה הוא כאשר אחד מהם מפסיק להסתכל

1. For every feature
   1. Count number of occurrences per video
   2. Sum lengths of occurrences per video

Calculate normalized values of (i), (ii). Normalize by length of interaction.

1. Correlation between couples of features. Check how to calculate correlation between binary variables. Cramers's V,
2. Uncertainty coefficient
3. Time window - Event frequency vs. time from event

מייצר היסטוגרמה שכל תא בה מייצג כמה פעמים קרה מספר אירועים מסויים בתוך חלון הזמן

1. Granger

לשים לב לכך שיש ל"נרמל" את הדאטא שיהיה סטציונרי אחרת לא ניתן לעשות גרנגר

איך סופרים ארועים? עבור כל אירוע שקרה ניקח את חלון הזמן שלנו בשניות,(ערך תחתון להתחלה,ערך עליון לסוף) ונסמן 1 עבור כל שניה בתוך החלון(אם חלק מהשניה בחלון)

לדבר עם עומר:

הוספצי מטריצת גרנגר

התיוגים הפעם שונים... לדוגמא verbal scaffolding-cp

זה עלול לגרור בעיות ניתוח, בנוסף יש שורות עם דאטא חסר, כתבתי פונקצייה למילוי דאטא