

# Social Rhythm and Online Courtship

Elizabeth Bruch<sup>1,2</sup>, Hanbo Sun<sup>1</sup>

<sup>1</sup>Complex Systems, <sup>2</sup>Department of Sociology, University of Michigan



## INTRODUCTION

**Goal:** Identify rhythms of human courtship and to what extent the patterns reflect the outcomes of courtship, biology, as well as social norms.

**Data:** 27 Million American users profiles, and their messaging and browsing history over 15 months from a popular dating website.

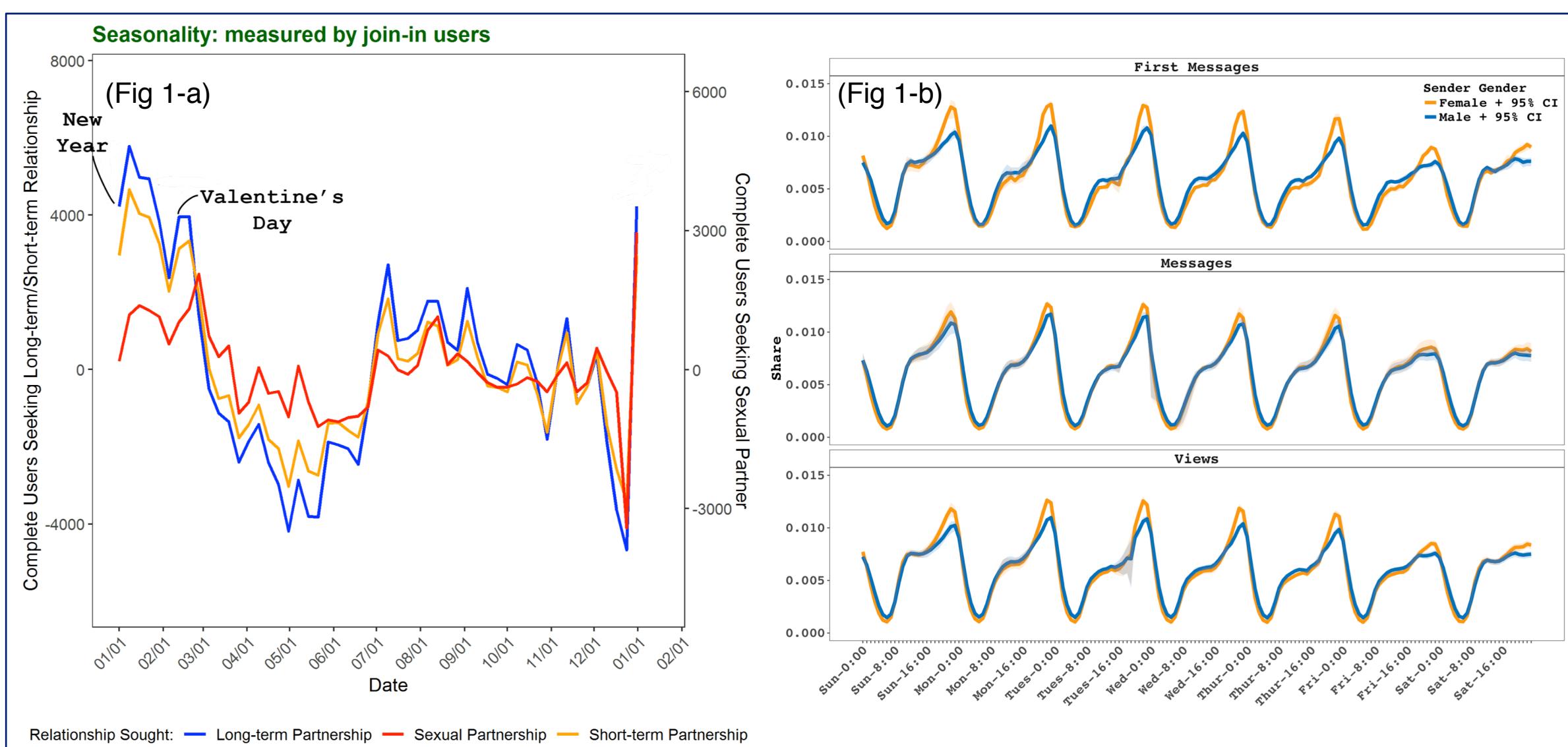
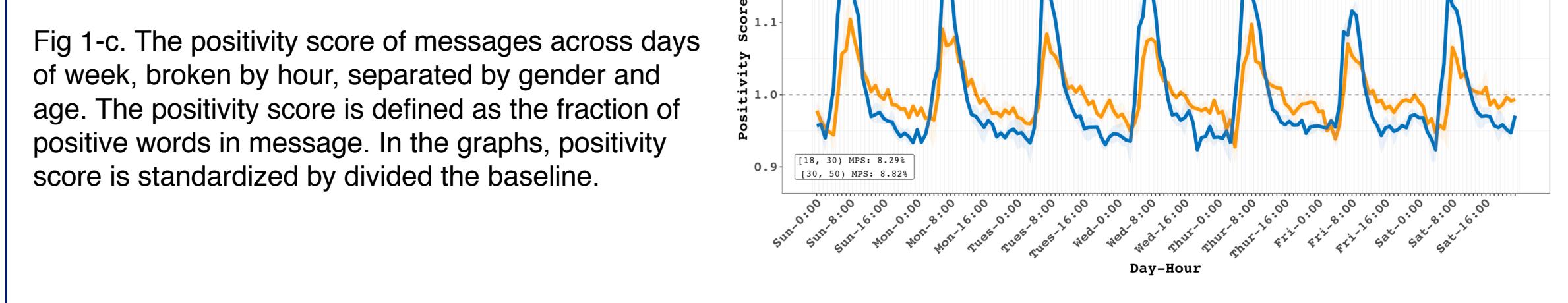


Fig 1-b. Number of initial messaging, overall messaging and browsing activities across week, broken by hours. Each curve is sum to 1 so the y-axis is the share of the activities in a given hour of a given day.

Fig 1-c. The positivity score of messages across days of week, broken by hour, separated by gender and age. The positivity score is defined as the fraction of positive words in message. In the graphs, positivity score is standardized by divided the baseline.



## METHODS

- **Seasonality analysis:**
  - Time series additive decomposition
  - Bootstrap confidence interval
- **Canonical rhythms recognition:** conversations clustering
  - Methods: K-means(++) , Mean Shift, Mini Batch K-Means, Spectral clustering, DBSCAN, Affinity Propagation, Birch, Gaussian Mixture Modeling
  - Overall metrics: number of messages, average waiting time
  - Asymmetry metrics: the difference of waiting time or message length between initiator and receiver
  - Acceleration metrics: waiting time  $\Delta T_n = T_n - T_{n-1}$ , message length  $\Delta W_n = W_n - W_{n-1}$ , the conversation is accelerating if  $\Delta T_n < 0$  or  $\Delta W_n > 0$ , where n is the index of non-redundant message in the conversation .
- **Association between rhythm and courtship outcomes**
  - Outcome: whether exchange contact information
  - Pleasure: the fraction of positive words in conversation

## RESULTS

- **Seasonality analysis:**
  - The majority of messaging and browsing activities takes place between 8pm and 12am.
  - The distribution of activities is consistent with the effects of sleep, circadian and work rhythms
  - People send the most negatively connotated messages in the late evening and the most positive messages when they wake up.
  - People (1) send more messages in late evening and Spring; (2) send longer messages in late evening on weekday, late morning on weekend and Summer; (3) reply quicker in Winter.
  - People living in bay area reply slower and sent messages less frequently, while people living in Florida are on the opposite.
  - ...
- **Canonical rhythms recognition**
  - 8 canonical rhythms of dyads interaction.
- **Association between rhythm and courtship outcomes**
  - Cluster 1 in 14% of conversations leads to a 1.8-3.2 times probability in successful courtship and associated with significantly more pleasant conversation. (Fig 4-a)
  - The conversation rhythms/patterns reflect social norm. (Fig 4-b, Fig 4-c)

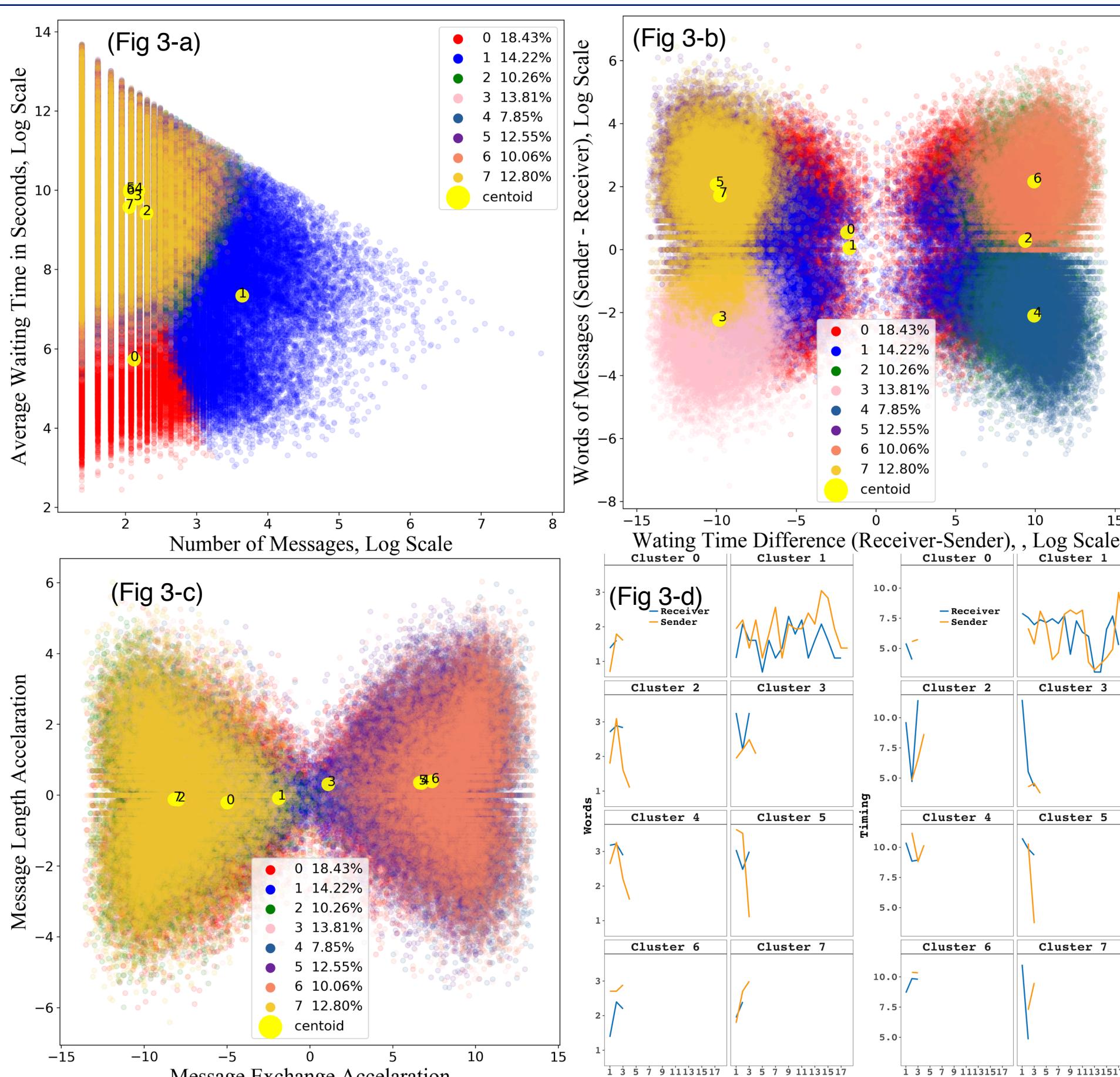
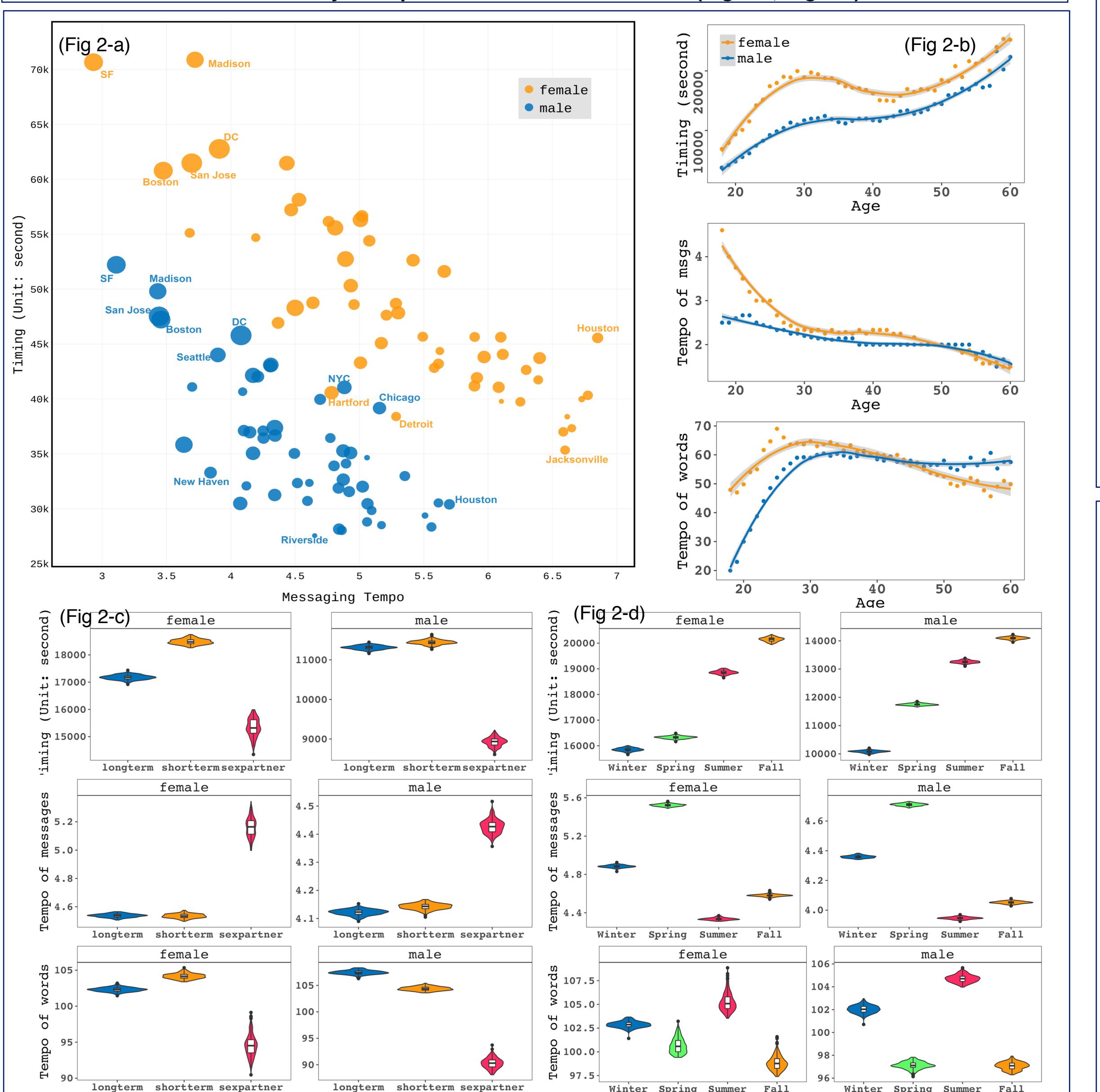


Fig 3-a,b,c. k-means++ clustering projection on variables' space. Within 9 clustering methods, k-means++ resulted in the best performance in terms of WCSS and Silhouette. The performances of mini-batch k-means and spectral clustering are comparable to k-means++, where they are consistent to suggest k = 8 or 9. Test on conversation in January and March separately, which results in identical clustering pattern. This set of graphs is variable space projection of 1.35M valid conversations in January. Fig 3-d. canonical rhythms of interaction, i.e., the central conversation of each cluster.

