When I Move, You Move: Coordination in Conversation

Gina M. Caucci,a Rick Dale,b & Roger J. Kreuza
University of Memphisa
University of California, Mercedb

Introduction

• A conversation is a complex, yet effortless joint action which requires coordination at some level for effective communication to occur (Clark & Brennan, 1991; Clark & Krych, 2000).
• Coordination requires timing.
• To explicate coordination requires an understanding of the foundation of conversational exchange: turn-taking (TT) (Benus, 2009; Bosch, Oostdijik & Boves, 2005; Sacks, Schegloff & Jefferson, 1974).
• TT here refers to the back-and-forth exchange of spoken information (Clark & Brennan, 1991).
• Researchers have found evidence that coordination explains this timing process explicitly using rate of turn (Schegloff, 2006; Stivers et al., 2009).
• Specifically, it has been proposed that an oscillator model of conversational coordination explains this timing process (e.g., Benus, 2009; Wilson & Wilson, 2005).
• Since TT is theoretically the timing of a dialog, understanding it is crucial for explaining the coordination or synchrony of conversational participants.
• Thus, the purpose of this study was to test this model explicitly using rate of turn-taking as the impetus for movement coordination.

Hypotheses

• Interlocutors will synchronize paralinguistic cues as a function of their rate of turn-taking.
• The more coordinated their rate of speech, the more they will coordinate movement.

Methods

Participants

• Videotaped interactions of 11 dyads engaging in conversation.
• Five minute speech segments were randomly selected from each dyad for analysis.

Procedure

• An estimated rate of TT was computed as a function of speech rate disparity between members of a dyad.
• Larger speech rates indicate more speech, longer turns and thus fewer turn changes throughout (see Figure 1.).
• Image sequences were created from the videos at a rate of 8 fps.
• Body movement was then measured using an image-differencing algorithm in MATLAB.
• Frame-by-frame movement was calculated by averaging pixel values across images.
• TT rate was then used to inform window size selection for the windowed cross-correlation analysis (Boker, Xu, Rotondo, & King, 2002) (see Table 1.).
• This method produces mean correlations between the short time windows from signal A & signal B.

Results

Table 1. Window and lag for each dyad.

<table>
<thead>
<tr>
<th>Dyad</th>
<th>TT rate</th>
<th>Window (sec)</th>
<th>Lag (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.18</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>0.37</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>1.08</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>0.15</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>0.49</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>0.20</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>7</td>
<td>0.35</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>0.87</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>1.03</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>1.15</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>0.90</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>

TT rate is represented as the difference in speech rate between the members of a dyad.

Figure 1. A comparison of two speech signals from two dyads.

Figure 2. Density plot from session 10 with the largest difference in speech rate $W_{max}$ of 6 seconds and $T_{max}$ of 4 seconds.

Figure 3. Density plot from session 6, which had one of the smallest differences in speech rate of the 11 dyads; $W_{max}$ is 3 seconds with $T_{max}$ of 2 seconds.

Figure 6. Density plot from session 1, which had the smallest difference in speech rate $W_{max}$ of 2 seconds and a $T_{max}$ of 2 seconds.

• Mean correlation vectors were submitted to a one-way ANOVA with window size as IVs and mean correlations as DVs.
• Results show a significant difference between windows $F = 14.02$ (df = 5, 5207), $p < .001$.

Discussion

• These results lend support to an oscillator model of turn-taking as it does appear that a precise timing mechanism inherent in turn-taking is generating rhythmically entrained nonverbal behavior.
• While preliminary, these data reveal an interesting area of exploration which focuses on social coordination without an examination of linguistic data.
• Future planned research includes an analysis of the current data using mixed dyads as well as a cross-modal measure of coordination.

References