PURPOSE

• To investigate the rhythmic structuring of fundamental frequency ($f_0$) and intensity in conversation interaction.
WORKING HYPOTHESES

• Rhythmic structure is commonly evident in spectral analyses of $f_0$ and Intensity data from conversational speech.

• Dominant cycles are frequently matched in period and phase across utterances, both across utterances by one participant, and across utterances by both participants (e.g., across “turns” and within episodes of simultaneous speech). We regard the latter situation as “interpersonal synchrony.”

• Cross-utterance continuity in rhythmic structure can be found in conversations between strangers and in different language communities.
CORPUSS

- Four 23 second extracts from 15 minute conversations between same-gendered and approximately same-aged strangers:
  - Two American English conversations recorded in Madison, Wisconsin (one between women and one between men), and
  - Two Swedish conversations recorded in Umeå (one between women and one between men).
**RECORDING AND ANALYSIS METHODS**

- Conversations are recorded using both lavalier condenser microphones and accelerometric “contact” microphones adhered to the throat.
- No topic is suggested; participants are asked to “get to know one another.”
- Both microphone signals from a passage at least 10 minutes into the conversations are digitized and prosodic parameters extracted in the CSpeechSP environment on PC microcomputer.
  - $f_0$ is analyzed using an autocorrelation-based algorithm operating on contact mike signals.
  - Intensity is analyzed using an RMS algorithm on condenser mike signals.
  - Both parameter extractions are carefully inspected for artifacts and hand-corrected as necessary
    - Erratic $f_0$ passages (e.g., due to vocal fry) are interpolated if brief (< 150 ms) and surrounded by reliable $f_0$, or deleted otherwise.
    - RMS data are edited to remove cross-talk and between-utterance energy.
- 23 second passages of prosodic data centering on a turn-exchange are down-sampled to 16 samples per second and exported to Systat environment.
- Successive 128 point (8 s) frames with 48 point (3 s) overlap are FFT analyzed with prior subtraction of frame mean, normalization for whole passage SD, and cosine edge tapering.
  - In addition, $f_0$ data are zero-padded (i.e., at the frame mean) to cross gaps in the data.
- Periodograms are 3-point smoothed to yield power spectra.
INTERPRETATIONS

• All individuals showed evidence of cyclicity in the spectra.

• We examined evidence for interpersonal synchrony by:
  1. Locating a strong period in intensity, f0, or both, for the person speaking at the outset of the passage (person “A”);
  2. Mapping the indicated period and phase onto person A’s prosodic data;
  3. Examining prosodic spectra of person “B” for evidence of cyclicity at a similar period;
  4. Mapping the indicated period and phase onto person B’s prosodic data;
  5. Checking for a graphic fit between
     a) A’s or B’s cyclic models and their individual prosodic data (e.g. stress peaks in f0),
     b) Cross-utterance continuity across A’s or B’s utterances, and
     c) Interpersonal synchrony across A’s and B’s cyclic models (sometimes seen by extension of A’s model, as illustrated in the Swedish Women with a dashed line).

• Interpersonal synchrony seems clear in three of the four samples, but less clear in the conversation between the American men.

• The American men nonetheless exhibited among the strongest tendencies towards maintenance of cyclicity across an individual speaker’s utterances.

• Future work will include more samples, other sampling rates to detect faster or slower cycles, and the use of cross-spectral analyses.
f₀ model of speaker A

dB model of speaker B

f₀ model of speaker B

dB model of speaker A

AMERICAN WOMEN
SWEDISH WOMEN