**Introduction**

Naturalistic stimuli drive similar neural responses across people. The synchronization of neural signals, calculated as a voxel-wise correlation in fMRI activity over time between subjects, can be driven by a number of engaging stimuli. Studies using audio and visual clips from movies posit that the strong inter-subject synchronization is driven by the level of engagement with the story a participant experiences (Naci, 2017). However, inter-subject synchrony has also been found between participants listening to classical music that lacked a language component (Abrams, 2013). In this case, consistent and reliable patterns of activity over time were seen in both cortical and subcortical brain areas previously shown to be involved in music processing indicating that there is more to synchrony than the presence of language information.

Music is different than other stimuli because it creates behavioural synchronicity. It is well known that people synchronize their movements as they clap or dance along to music (). It is possible that the mechanisms involved in allowing people to predict the beat in music will show high inter-subject synchronization using imaging techniques like fMRI. Music is also different than other stimuli in the way that it is remembered over long periods of time. There is evidence to show that older adults with neurodegenerative disorders remember music from their youth even after other semantic memories have been forgotten (). Traditionally, the method for exploring memory for music has involved a comparison between the BOLD activation levels during unknown and well known music (). One downside to this approach, is that music unfolds over time and a BOLD contrast collapses over the time dimension that is unique to music. An inter-subject synchronization approach may be a more sensitive method to use to understand why memory for music is remembered differently than other semantic stimuli.

No studies to date have explored how inter-subject synchrony varies as a function of the presence of language and music. In this study we explore shared neural representations underlying the processing of naturalistic stimuli that vary in the presence of music and language and how that representation changes as people become more familiar with the stimuli.

If rhythm and music are added to the spoken word stories that create high ISC, do we see a change in ISC. Specifically, does music drive ISC as strongly as, or more strongly than spoken word alone.

See decrease in synchrony with familiarity – shown in EEG and in fMRI work. This has been shown with stimuli that contained words (either words themselves, or movies). What happens when music is involved?

You synchronize to naturalistic stimuli whether you like it or not.

**Methods**

**Participants**

Twenty-six neurologically healthy participants (14 female) aged 18-39 (mean=24) were recruited via posters and word of mouth at The University of Western Ontario. *Include relevant musical demographic information based on what goes into the analysis.*

**Testing procedure**

Participants completed two functional MRI scans that were separated by a stimulus training period (14-29 days; mean = 19 days). During both scans, participants passively listened to the stimuli (described below). During the training period, participants listened to the stimuli via an online player (designed in-lab) that tracked the number of times each stimulus was played. To ensure participants were listening, the player presented a simple question about the stimulus (e.g. “*were there lyrics present in the previous song?”*) at randomly between stimuli. Participants also came to the lab between scans four times. In each of these sessions, participants listened to the stimuli in lab and completed a serious of behavioural tasks (described below).

**Stimuli.** Eight different auditory stimuli were created from songs written and recorded by a lab member between *year1-year2* in Cambridge, UK. These songs were chosen to lower the likelihood of participants being familiar with the stimuli. Two songs were kept whole (vocals & instruments), two songs had the vocals removed leaving just the instruments, two songs had the instruments removed leaving a single a capella voice, and the lyrics of two songs were recorded in-lab as spoken word (no music). The voice in all stimuli was the same. All stimuli were extended to 5 minutes long (using *Audacity*) by repeating a chorus or verse where needed. During the training period, participants listened to half of the stimuli via the online player (4 songs, one of each type). The training sets of stimuli were counterbalanced across participants.

**Behavioural tasks.** During the training period, participants came into the lab four times. Each session lasted less than one hour. Participants listened to the stimuli in-lab and completed 2-3 of the following behavioural tasks in each session.

To test whether participants were learning their training stimulus, two tests designed in-lab were used. The first, was a lyric modification task that presented participants with pairs of lyrics. Each pair consisted of a lyric taken directly from their stimulus training group and a modified version of the same lyric. Participants indicated which lyric was the correct lyric. The lyric pairs were tested for their validity before being included in this study. Before the first scan session, participants were presented with the entire set of 25 lyric pairs to obtain a baseline measurement. As participants were not familiar with the stimuli these lyrics were taken from, they were asked to indicate which lyric they believed was most likely to come from a song. In each behavioural session, participants responded to a subset of 10 lyric pairs to track learning progress. After the second scan session participants completed the full set of 25 lyric pairs again. The second test of familiarity was a melody recognition task. After the second scan, participants heard pairs of 2 sec clips taken from the stimuli. These clips did not contain any lyrics. One clip was taken from the stimulus training set, the other clip was from a stimulus the participant did not train on. Participants were asked to indicate which clip was most familiar to them.

Participants completed a questionnaire regarding musical abilities and training as well as a test of melodic memory and a test of beat perception taken from the Goldsmith’s Musical Sophistication Index (CITE). Finally, participants completed a musical association form where they described what each of the trained stimuli reminded them of (memories, other songs, etc.) and a lyric orientation questionnaire that measured to what degree the participant focuses on the lyrics in a song over the melodic content (D. VUVAN).

**~~Preference ratings.~~** ~~In each lab session and after the second scan, participants rated on a scale of 1-7 how much the liked the songs in their stimulus training group allowing us to track how preference changed as familiarity increased.~~

**fMRI acquisition and analyses**

Imaging was conducted at the Robarts Research Institute on a Siemens Magnetom 7 Tesla scanner with a 32-channel head coil. Functional scans were acquired with 54 slices per volume (TR = 1.25 s; TE = 20 ms; flip angle = 35°; FOV = 220 x 220 mm; voxel size 2.5 mm3). Between functional runs within the first session only, a whole-head anatomical scan was acquired (TR = 6s; TE = 2.69 ms; FOV = 240 x 240 mm; voxel size = 0.75 mm3; 208 slices). Eight functional runs, each lasting five minutes (the length of the stimuli) were collected in both the first and second scans. The order of the eight songs was randomized in each scanning session for each participant.

Data were processed using SPM12 and automatic analysis software (AA; [www.cusacklab.org](http://www.cusacklab.org)). Data were corrected for motion and normalized to a template brain. Regressors were calculated to account for artifacts in white matter and cerebrospinal fluid. Smoothing was done with a Gaussian kernel of 8 mm FWHM (Peigneux et al., 2006). These normalization parameters were then applied to all echoplanar images.

Intersubject correlation (ISC) was calculated for each stimulus by extracting the timecourse of every voxel and taking a correlation between each subject’s voxel timecourses and the mean timecourses from every other participant (leave-one-out method). This created a map of correlation values for each participant indicating their degree of correlation to the group. Performing a t-test on each voxel produced a spatial map of areas where the group was highly synchronized while listening to that stimulus.

The same ISC calculation was performed in specific ROIs. The Yeo 7 network parcellation (CITE) was used to identify frontoparietal areas. Three other ROIs were used as defined by neurosynth.org: motor (from 2565 studies), basal ganglia (from 438 studies), and auditory (from 1252 studies).

**Results**

**Behavioural Results**

Over the course of the training period, participants listened to the stimuli on average 12.8 times (SD = 4.6). The familiarity tests indicate that behaviourally, participants learned the stimuli they trained on. A 2(session: 1 & 2) x 3(type: spoken, a capella, whole) ANOVA was run on the lyric modification results. There was a main effect of session (*F*(1,138)=159.2, *p*<0.001) with results increasing from an average of 36% correct (SD=13.7) before the first scan, to an average of 82% correct (SD=9.8) after the second scan (see Figure 1). There was no main effect of stimulus type (*F*(2,138)=3.0, *p*=0.05). The melody recognition results collected after the second scan were at ceiling (mean=92%, SD=6.4) (see Figure 1). Four participants were excluded from the rest of the data analysis, because their average score across both memory tests did not reach 70%.

**fMRI results**

**Discussion**

Bring in consciousness piece – about how ISC may be a measure for understanding how we and others perceive our world.

Future research – older adults.