1. Preparation List
2. Main Goals
3. Design
    1) Stimuli
    2) Subjects Selection
    3) Experiment Design
    4) Behavior Test
    5) fMRI scanning
        Structure image 1
            Parameter
        Session 1(Reading)
            Parameter
            Instruction
            Presentation
            Posttest
        Structure image 2
        * Session 2(Listening)
            Parameter
            Instruction
            Presentation
            Posttest
        Structure image 3
        Session 3(Speaking)
            Parameter
            Instruction
            Presentation
    6) Counterbalanced Matrix
    7) Time estimation
    8) Expense estimation
4. Data preprocessing
    - Data Format Transformation
    - Single-Subject Analysis
        - Correct Head movement
        - Surface reconstrcution
        - Segmentation
        - Slicing Time
        - Registration
        - Alignment
        - Normallize
        - Smooth
5. Post processing
    - SRM
6. Conclusion & Significance

1. Change Record : First Version finished on 19/6/2018, second edited on 15/7/2018. Edited by Ruiqing Zhang
   Check List :
   - ❑ IRB
   - ❑ Informed Consent (English/Chinese verision)
   - ❑ Participant fee (30~40 RMB for Behavior test + 100 RMB /per hour in MRI scanning)
   - ❑ Scanning cost(1000RMB/per hour)
   - ❑ Earplugs & CDs

2. Main Goals

1) Differentiating

This study aims to establish a neural base of Chinese-English learners with different proficiency levels comprehending English in three modules and quantifying L2 ability from text-based test to neural-based test.

2) Logic

Language proficiency level is correlated with the functions of language neural network. For the low-level L2 learner, language-network is not well-suited and usually display divergent activation profile in fronto-tempo network and redundant exploitness. While for high proficiency level, information is transferred efficiently among fronto-tempo network. Behavioral comprehension of narrative stories may not exist differences, while the efficiency of language-network should have different activation profiles.

Hypothesis 1: Phoneme discrimination is related to listening ability. Behaviorally, a significant difference should be expected among L2 learners with graded proficiency. Then use phoneme discrimination ability as predictor, the brain mechanism explained the difference would be derived and mostly anticipatedly, the superior temporal cortex.

Hypothesis 2:

3. Design

★ Stimuli :

- ❑ Contents: the samples selected need to be understood by each subject especially the subject with low proficiency level.
  - ❑ fiction: Alice's Adventures in Wonderland, 2152 word
  - ❑ non-fiction: how many different kinds of emotions are there? 1949 word(selected from frontier for young minds)
- ❑ Speed: 0.7\ 0.8\ 0.9\ 1\ 1.1\ 1.2\ 1.3\ 1.4\ 1.5 (tempo changer)
- ❑ Audio materials: recorded by two native speakers(male voice and female voice), normal speed
- ❑ Speaking materials
- ❑ Subjects : APP users, gender balanced, N=72
  - ■ Native Speaker group : fluent speakers, normal hearing, n=18
  - ■ Chinese-English group : App Palcement Test(level1~level8)
    - low proficiency level group(15 people): App Placement test 1~3
    - Intermediate proficiency level group(15 people): App Placement test 4~5
    - High proficiency level group(15 people): App Placement test 6~8

★ Experiment Design

The subjects will fill out several detailed questionnaires to assess their language background and general L2 abilities.

- ➔ Behavior test :
  - ◆ Language Experience Questionnaire
  - ◆ Edinburgh Handness Questionnaire
  - ◆ Proficiency Questionnaire(LEAP-Q)
  - ◆ Phoneme Discrimination---related to listening ability
  - ◆ PictureNaming test---related to speech ability
  - ◆ Picture Vocabulary Size Test(PVST) ---related to reading ability
  - ◆ Reading Speed Test(http://www.freereadingtest.com/)---adjust the presentation speed in RSVP
  - ◆ Speech Assessment: record 2 min describing activity in one day using Praat software.

- ➔ fMRI scanning:
  - ◆ Preparation:
    - ❑ Do you have metal materials in your body, such as in the teeth, sugery inplantation, plastic sugery?
    - ❑ Do you wear any metal ornaments like rings, earring, necklace, hair clip, metal button or zipper, coins, watch, RF/IC card?
    - ❑ Do you have claustrophobia?
    - ❑ Ask for diopters if the subject is nearsighted and put on the metal-free glasses
    - ❑ Wear Earplugs
    - ❑ Digital number coding instead of Subject's name or initials
    - ❑ Compile the experiment scripts
    - ❑ Check earphone in both ears and volume
    - ❑ Adjust the screen position
    - ❑ Check the speaker and sound recorder.

  - ◆ fMRI Experiment Procedure:

The whole process would be like this:

First 6 min structure image scanning > Session 1 (Reading/ Listening/ Speaking) > post test > rest time > Second 6 min structure image scanning > Session 2 (Listening/ Reading/ Speaking) > post test > rest time > Third 6 min structure image scanning > Session 3 (Speaking/ Reading/ Listening) > post test

Instruction: Before step into scanner, let the participant know the whole procedure and time, and stay still as possible as they can once they find a comfortable position to avoid head movement during scanning. Send the subjects into the scanner (operate by the technician) and adjust the mirror to make sure they have a good vision of the projective screen. Play test sound and adjust the volume to a comfortable level.

- Scanner : Siemens MAGNETOM Prisma fit 3T
- fieldmap
- Structure image:

  Key Instruction: stay still as possible as you can while scanning in order to get a good-quality image. The structure image will cost about 6 minutes.

  Parameters: functional data will be collected on a 3T Siemens Prisma scanner with 32-channel head coil. the funtional images are obtained with an interleaved

single-band EPI sequence(TR=2530ms, echo time =2.34ms, flip angle(FA) =7°, slice thickness= 1mm, slice number =192, sagittal direction, Field Of View (FOV) = 258 mm²) T1* structural images are acquired with an MPRAGE sequence.

- ❖ Functional image : first, 10s dummy scan

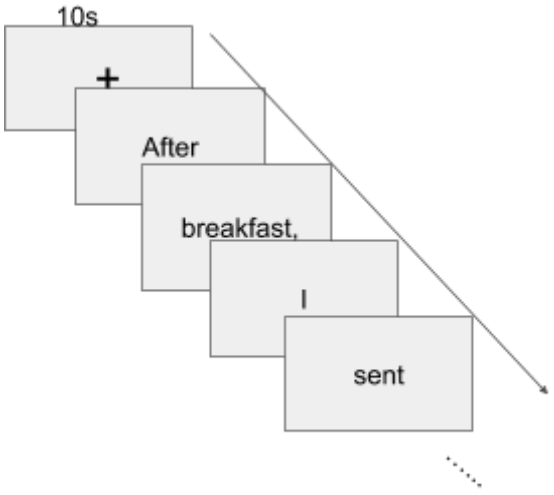
Sequence parameters :

Sequence	Duration	Volume	TR	TE	FA(flip angle)	slice thickness	slice number	Series	Orientation	FOV	accelerator factor	in-plane Resolution
T1	6 min	1	2530 ms	2.34ms	7°	1mm	192	Interleaved	Sagittal	258mm	single-shot	1*1mm²
listening-MB band	25min	1500	1000 ms	0.03s	62°	2.5mm	60	interleaved	Sagittal	100mm	4	2.5*2.5 mm²
Speaking-MB band	>24min	1500	1000 ms	0.03s	62°	2.5mm	60	interleaved	Sagittal	100mm	4	2.5*2.5 mm²

- ➢ Reading:

Each subject will go through a quick reading speed test (RST for short) at first and specific RST for specific difficulty. After collecting subjects' reading speed, the presentation rate of the reading samples on screen will adaptively matched. The procedure will start from an introduction page with a cue indicating session module in this case" " then the RSVP paradigm presenting single word.

In each run of reading session, the materials will be presented by a fixation, then followed word in rapid series(duration is depend on the reading speed assessed before fMRI scanning) till a story is over and wait for 10s blank screen and another story begins. The stories are played in randomized sequence.



Reading speed:
generally, the native speaker is over 200 wpm =300ms /per word
high level PL L2 learner can get 140~160 wpm = 429 ms/per word ~ 375ms /per word
inter level PL L2 learner can get 70~120 wpm = 857 ms/ per word ~500ms /per word
low level PL L2 learner can get 60 ~ 70 wpm = 1000ms/per word ~ 857 ms/per word

➢ Listening

The procedure will start from an introduction page followed by a cue indicating session module in this case a"  ", then a 6-min structure image is obtained. Then functional image collection start from audio playing with comprehension questions following afterwards.

➢ Speaking

The procedure will start from an introduction page followed by a cue indicating session module in this case a"  ", then a 6-min structure image is obtained. Same materials as in reading session except while subject is reading, they should articulate silently.

◆ Counterbalanced Matrix

Each group should balance sessions(repeated twice in each group):

session1	session2	session3
session1	session3	session2
session2	session3	session1
session2	session1	session3
session3	session1	session2
session3	session2	session1

4. Data Preprocessing

fMRI data processing uses FreeSurfer 6.0 and AFNI.
workflow:https://surfer.nmr.mgh.harvard.edu/fswiki/FsFastWorkFlows

- Data format transform and organization

The raw data obtained from scanning is DICOM, which cannot be used for processing. Use function dcmunpack or MRIConvert unpack Siemens DICOM files into 4D files(mgz, nii ...).
Check data: volumes number, slices, sessions hierarchy..etc

1) convert data format:

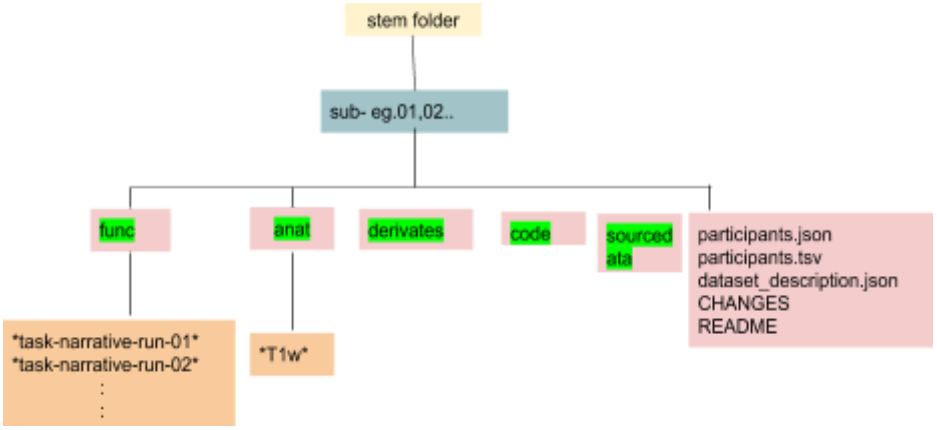
```
dcmunpack -src pilot_data \
   -trg mri \
   #-run 1 subID mgz 000\ #localizer
   -run 2 subID mgz 001 \ #t1
   -trg bold \
   -run 3 subID mgz 002 \ #run1
   -run 4 subID mgz 003 \ #run2
```

2) organize data :

Project folder is the session folder: Reading, Speaking, Listening

Then secondary path is subject folder(001,002,003...) and a txt file(subject ID list)

The subject folder contains functional folder(BOLD), anatomy folder(T1),session info (generated by last unpack step), txt file with subject name (convenient to the iteration of data analysis)



3) *Paradigm file ? is this step necessary in this experiment!? ver5 requires 4-column , over the forth column will be ignored.

Paradigm txt file structure(example)

Column1	Column2	Column3	Column4	Column5
onset time(s)	condition	duration	Weight	Name
(stimuli time - trigger time) *discard dummy scan time	numeric coding			
0.0000	1	6	1	face
6.0000	0	1	0	fixation
7.0000	3	6	1	house
13.0000	2	:	:	:
:	:	:	:	:
:	:			

* There are the same number paradigm files as the number of runs, in this experiment we will have three paradigm files created and each file need to be put in the same folder as their specific functional data and name of the files in each folder are the same.

- Single-Subject Analysis

B0 distortion correction: Freesurfer epidewarp.fsl

Pipeline:

First, check the data quality and quick view of the raw BOLD signal change in afni software to see if there are some apparent head movement need to correct.

Then extract the anatomical cortical surface and registered to the template. Following head movement correction and slice timing correction, the functional images need to be registered to their anatomical template. A gray matter mask is used and normalize the registered functional data to a common space. At last, a 5 FWHM kernel is used to smooth the images.

vi prepro.sh

Scripts running on freesurfer using preproc and seperate stages:

surface-based anatomical analysis:

set subject folder: export SUBJECTS_DIR=${PWD}#${PWD}="~/Desktop/pilot_data/"
subID={001,002...}
all procedure at the same time: recon-all -s subID -all
or in 3 seperate procedures: -autorecon1, -autorecon2, -autorecon3 #this way aim to check the skull striping, segmentation etc.
check recon-all result:
tkmedit subID T1.mgz -aux orig.mgz #contro-1,2 swithing the data to inspect the intensity
tkmedit bert brainmask.mgz -aux T1.mgz #check the skullstriping
tkmedit bert wm.mgz -aux brainmask.mgz #check the segmentation
brainmask.mgz is the file after skullstrip, T1.mgz and orig.mgz are the files transformed(orig uses the native space).

freesurfer to SUMA: resamples the surfaces to standard topologies (with different resolutions) using MapIcosehedron, aligns surfaces to a reference functional volume, and merges left and right hemispheres into single surface files.
eg : @SUMA_Make_Spec_FS -NIFTI -fspath sub001/anatomy -sid 1

functinal data preprocessing:
No.1

```
preproc-sess -s sess01 -fsd bold -stc up -surface fsaverage lhrh -mni305 -fwhm 5 -per-run
```

note: make sure the subjectname file contains only the subject name under processing, or the "if: expression error" would happen. -mni305 is for subcortical registration. -fwhm can be set to 0 for hyperalignmen analysis.

- Registration Template Creation

```
mri_info template.nii.gz
mri_info f.nii.gz
```

- Correct Head Movement
  cross-session head correction

  fmcpr.nii.gz – this is the motion corrected functional data. It has the same size and dimension as the raw functional data.
  fmcpr.mcdat - text file of the amount of motion at each time point. This is important for Quality Assurance (see below).
  mcprextreg - text file of the motion correction parameters assembled into an orthogonalized matrix that can be used as nuisance regressors
  Verify that fmcpr.nii.gz has the same dimensions as f.nii.gz using mri_info (see example under section 4.1).

- Slicing Time Correction

  Verify that fmcpr.up.nii.gz has the same dimensions as f.nii.gz using mri_info

- Functional-Anatomical Registration

  register.dof6.lta - is a text file that contains the registration matrix.
  register.dof6.mincost - is a text file that contains a measure of the quality of the registration.

- Mask Creation

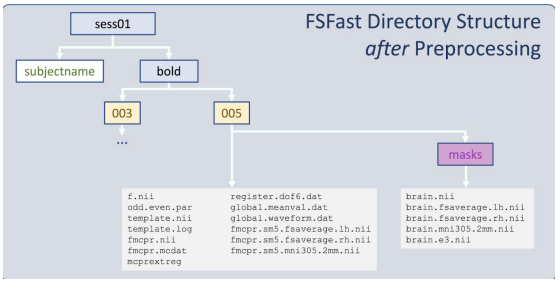
```
freeview -v template.nii.gz masks/brain.e3.nii.gz:colormap=heat -viewport coronal
```

- Normallization

  'text file global.waveform.dat'

- Resampling Time Series
- Smoothing

```
mri_info --dim fmcpr.up.sm5.fsaverage.lh.nii.gz
mri_info --res fmcpr.up.sm5.fsaverage.lh.nii.gz
```

After preprocessing folder structure: 

- Quality Assurance
  - Check raw functional images individually: artifacts and distortions, align the images by hand, apparent head motion in a movie.
  - Check transformations: be careful with the raw data transformation and functional data to anatomical data transformation(well aligned?)

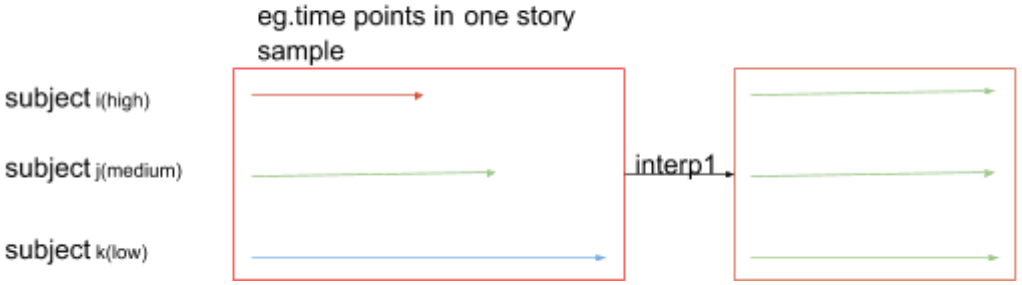
No.2:

surface preprocessing using afni_proc.py scripts(Desktop/RSA/preproc.sh: bash preproc.sh) *note: this would be done in seperate runs in each subject.

anatomical data and surface data use the brain.nii(copy anatomical dset) and *_SurfVol.nii and spec file [1,2]_?h.spec under subjects dir/anatomy/SUMA/

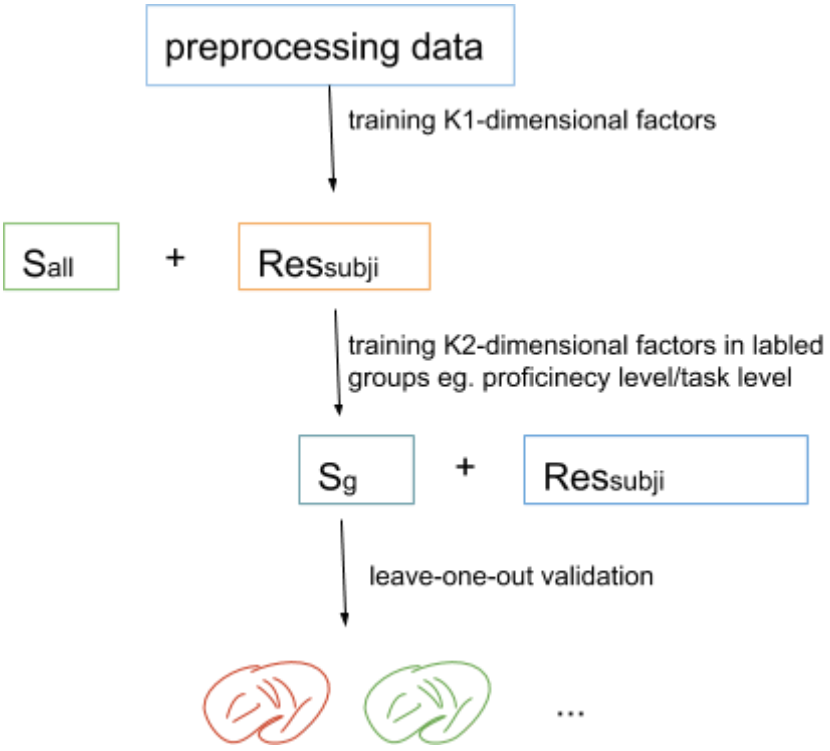
5.Postprocessing（http://brainiak.org/docs/index.html）

- Time scaling of functional data in each story across subjects to acquire equal time points：interp1 in matlab (Temporal scaling of neural responses to compressed and dilated natural speech. Y. Lerner, C. J. Honey, M. Katkov, and U. Hasson)



- Combine the functional data of all story samples in the same task level(easy, intermediate, difficult) individually：3dDeconvolve
- Differentiate different groups:
  - ❏ Shared response model: http://brainiak.org/docs/brainiak.funcalign.html
    - ❏ Step 1: train original data $X_{1:m}$ of all subjects to learn a k_1 dimensional shared response S_{all} and specific subject base $W_{1:m}$. Mapping the S_{all} in subject i topography and subtract from the original data to form a residual response X_i - W_i
    - ❏ Step 2: labeled group model training and leave-one-out validation. Within group, a K_2 dimensional shared response will be trained to obtain the shared

response S_{g1} S_{g2} S_{g3} and respectively subject base $W_{1:m(g1)}$ $W_{1:m(g2)}$ $W_{1:m(g3)}$ and test on the classier on held-out subject.



6. Conclusion & Significance

An expected cortical indicator of proficiency level will be obtained and different representative patterns in L2 learners with different proficiency level will be discussed.