The neural substrates of speech perception in Mandarin and English

Sophie Scott

Stuart Rosen

Yui Yu Meng, Jane Warren

Richard Wise

Mandarin (Putonghua, ‘the common language’)

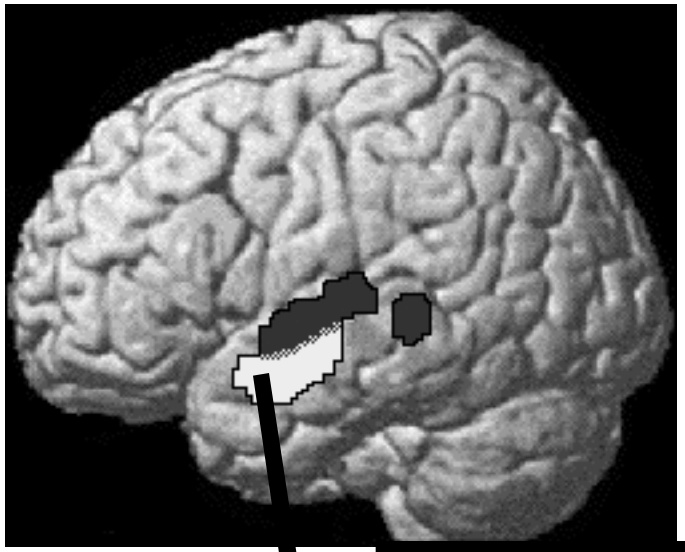
- The most widely spoken language in the world.
- It is characterized by a strict, simple syllable structure (CVC where the final C must be n or N).
- The use of tone as a lexical contrast.
- An ‘aspect’ syntax (e.g. no marking of tense or case).

English

- English is the most widely spoken Germanic language.
- It has a complex syllable structure
  (CCCVCCCC is a legal syllable - structure within this governed approximately by sonority)
- It has relatively simple ‘tense’ syntax.
- A wide vocabulary.

Neural basis of speech perception

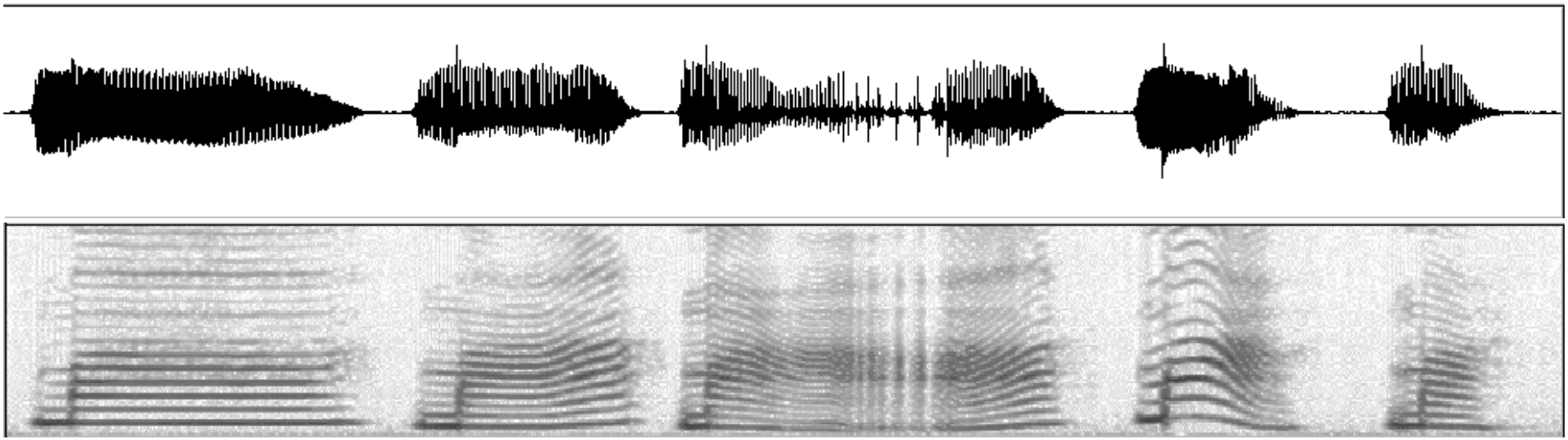
- In a previous study we showed that with appropriate controls for the acoustic structure, intelligible (English) speech is processed in the left STS, and dynamic pitch is processed in the right STG/STS Scott et al, 2000



Regions in yellow respond only to stimuli which are intelligible

5 Mandarin tones - ‘ma’

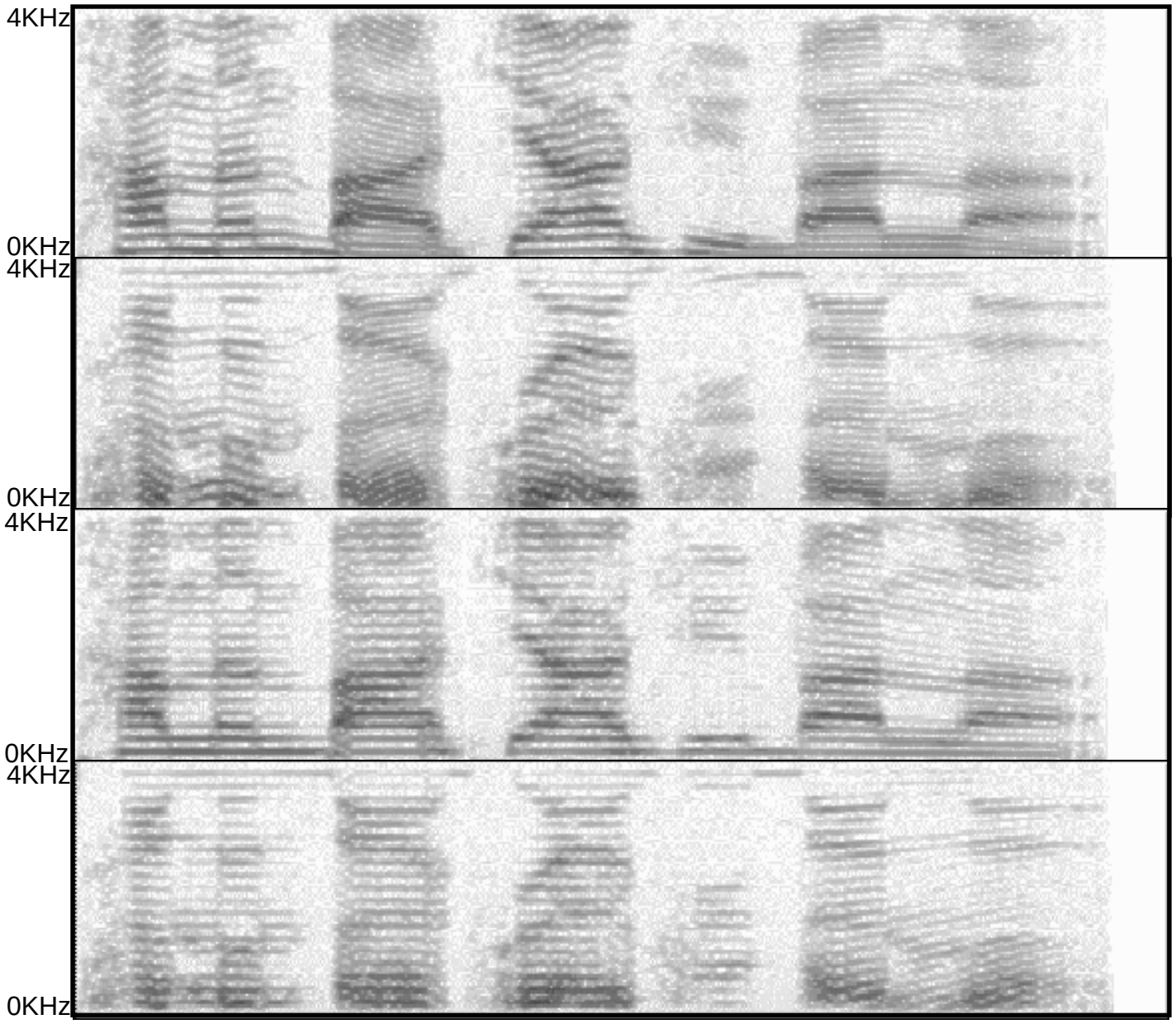
level rising falling - rising falling neutral



Design

- Mandarin translations of the BKB sentences were used.
- 5 conditions: normal Mandarin (small pitch shifts introduced, low pass filtered at 4KHz), spectrally rotated (at 2HKz) Mandarin, ‘detoned’ Mandarin (with a normal sentence intonation but no lexical tone), rotated de-toned Mandarin and rotated noise vocoded Mandarin.
- 6 RH male native Mandarin speakers.
- PET scanning, passive speech perception.

Spectrograms

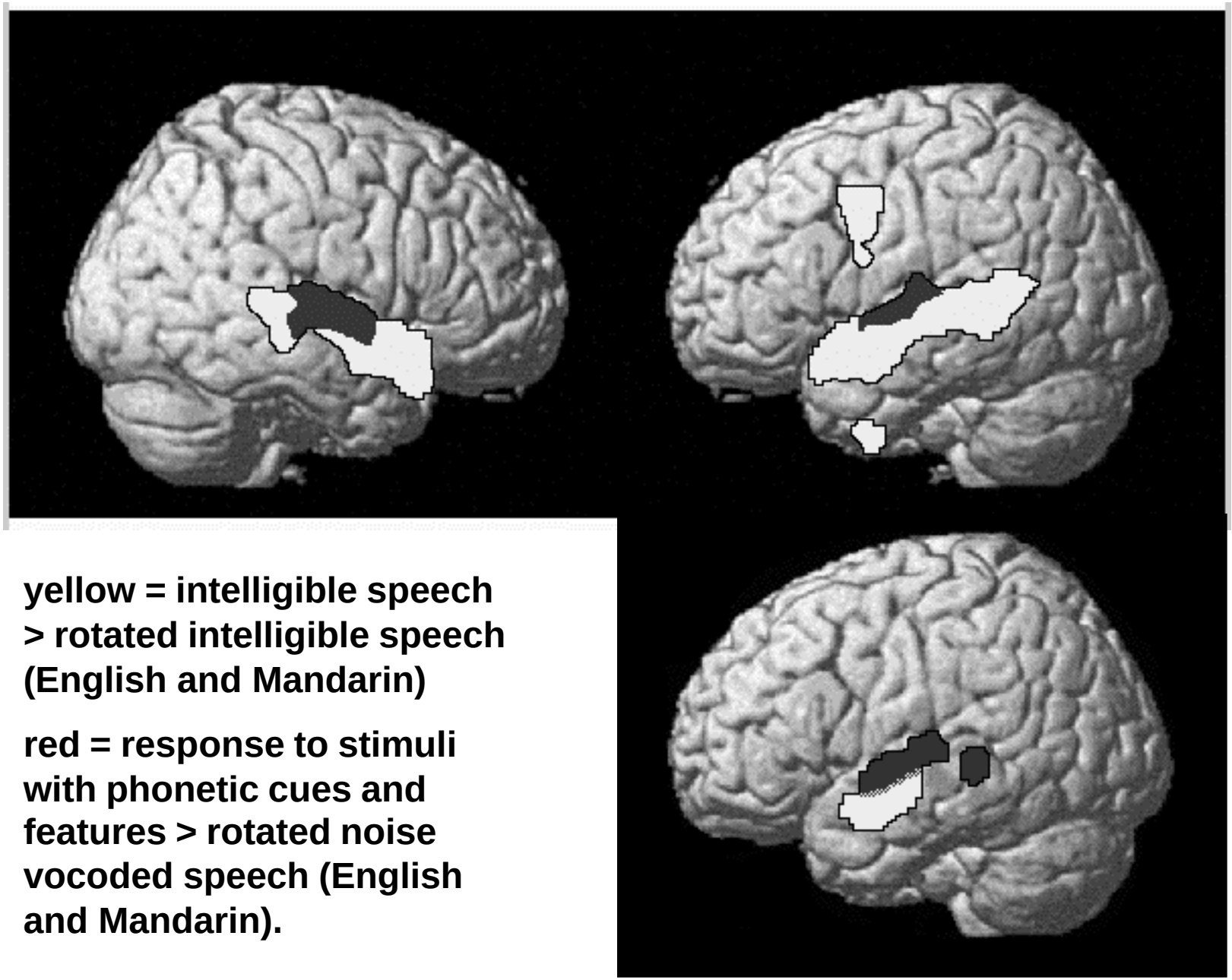


Low pass filtered Mandarin

Spectrally rotated Mandarin

‘Detoned’ Mandarin

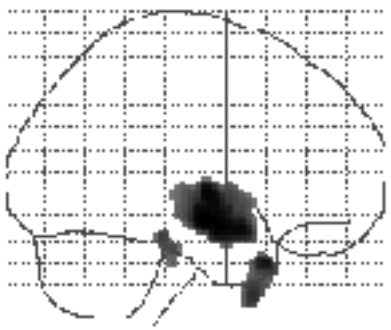
Rotated ‘Detoned’ Mandarin

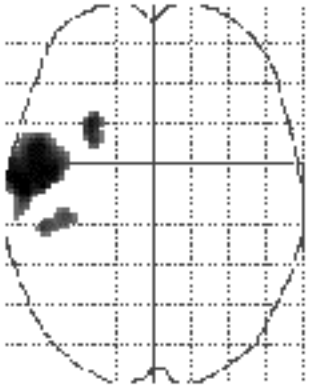


yellow = intelligible speech > rotated intelligible speech (English and Mandarin)

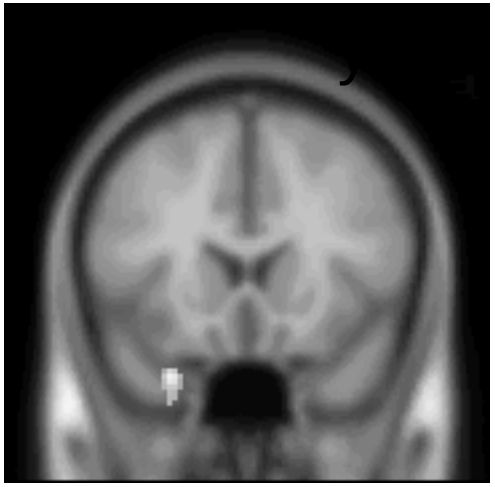
red = response to stimuli with phonetic cues and features > rotated noise vocoded speech (English and Mandarin).

Regions common to Mandarin and English

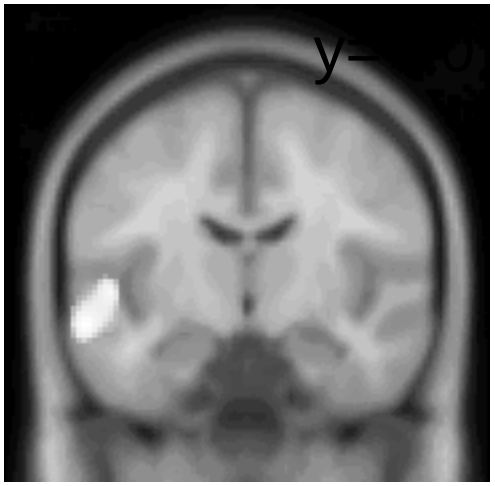




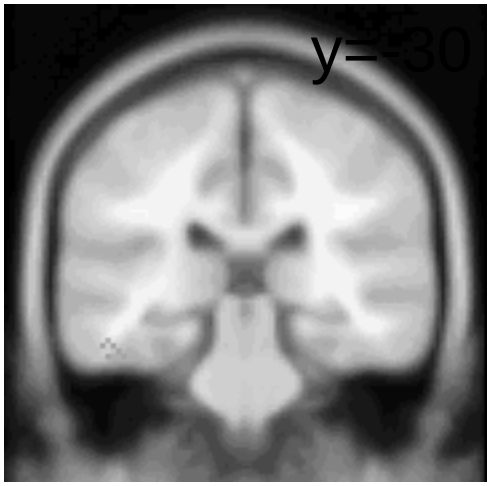
Random effects comparison, p<0.0001, k>40)



Z = 4.67

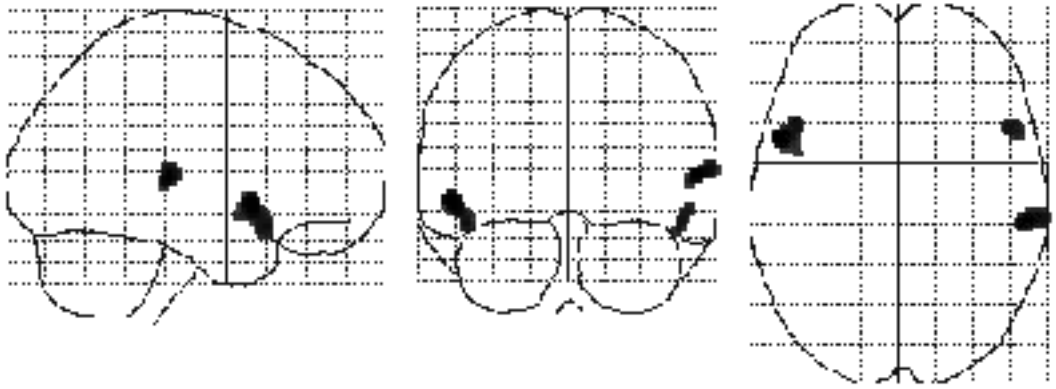


Z = 5.35

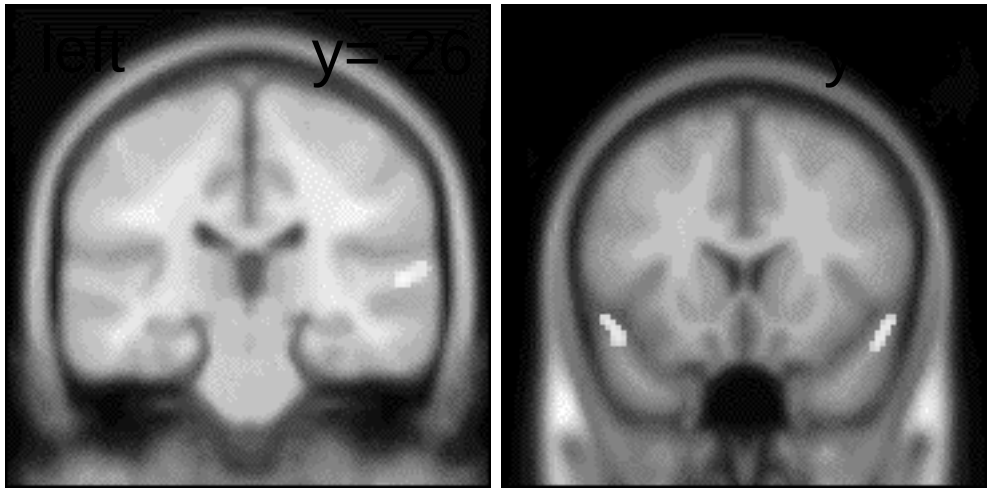


Z = 4.22

Regions more activated in Mandarin than English listeners



Random effects comparison, p<0.0001, k>40)



Z = 4.67

Z = 4.91, 4.03

Results

- Both Mandarin and English listeners recruit left anterior STS, temporal pole and fusiform gyrus.
- Mandarin speakers recruit, in addition, left and right STG and right posterior STS when processing speech.
- There were no regions of increased activation associated with English.

Conclusions - common systems

- Consistent with previous findings, both English and Mandarin speakers recruit a left hemisphere temporal lobe system when passively listening to speech.
- Speech specific activations were seen in the anterior STS, which has been implicated in the mapping of sound onto meaning, and in ventral regions associated with semantic representations.

Conclusions - language specific responses

- Mandarin speakers showed right posterior STS and bilateral anterior STG responses for both normal and detoned mandarin.
- Is this due to the lexical use of pitch variation? (pitch dynamics are associated with the right STS/STG Zatorre and Belin, 2001, Patterson et al, 2002)
- Is this due to a greater reliance on pragmatics? (also associated with right hemisphere functions)
- Is the greater left STG involved in integration of such information?