Contribution of the Visual Word Form Area to speech processing: When and how?

Chotiga PATTAMADILOK

chtiga.pattamadilok@lpl-aix.fr

Laboratoire Parole et Langage, Aix-Marseille University
Visual Word Form Area & Reading

Centred on the occipito-temporal sulcus at the transition from the occipital to the temporal lobe.

Its functional role emerges with reading acquisition. The activation level depends on reading expertise (Brem et al. 2010; Dehaene et al., 2010).

Play a critical role in reading (Duncan et al. 2009; Gaillard et al., 2006)

Reproducible activity across writing systems and individuals (Bolger, Perfetti & Schneider, 2005; Rueckl et al., 2015)
Neural model of reading

Access to pronunciation and articulation

Access to meaning

Visual Word Form Area

Visual inputs

Dehaene 2009

Thiebaut de Schotten et al., 2014

Anatomical connections

- Arcuate long segment
- Arcuate posterior segment
- Arcuate anterior segment
- Inferior fronto-occipital fasciculus
- Inferior longitudinal fasciculus

Dehaene 2009
VWFA responses during speech processing

- Speech processing tasks that **require an activation of orthographic knowledge**. E.g., do spoken words share the same spelling (Booth et al, 2002) or contain a specific letter (Ludersdorfer et al., 2015)

Spoken input | Target letter
---|---
/faːzə/ (PHASE) | « p »
/faːzə/ (PHASE) | « a »
/tiska/ (TISKA) | « i »

Ludersdorfer et al., 2015
• Speech processing tasks that **do not** require an activation of orthographic knowledge

Rhyme judgment vs. melody judgment (Yoncheva et al., 2009)

Selectively attending to speech, relative to selectively attending to melody, leads to increased activity in the VWFA

Yoncheva et al., 2009

“cross-modal sensory suppression” phenomenon
• Speech processing tasks that do not require an activation of orthographic knowledge

Comparisons between orthographic, semantic and tone processing tasks.

Auditory lexical decision
The activation level depends on reading expertise

Ludersdorfer, 2016

Dehaene et al. 2010
What happens in more natural speech processing situations?
- Passive sentence listening

- No significant VWFA activity
- The level of VWFA activity did not vary with reading expertise

Dehaene et al., 2010

Sentences instead of single words which make sound to spelling conversion less likely?
Passive situation?
“Rest” rather than non-language auditory baseline?
Other reasons?
**VWFA involvement in spoken sentence processing**


**Aim:** Identify the circumstances leading to VWFA responses to spoken sentences

**Participants:** 24 adults, right handed, skilled readers.

**Stimuli:** Spoken sentences

**Task demands:** Comprehension (false statement detection) vs. Perceptual task (one-back)

**Signal quality:** Clear vs. degraded listening condition (multi-speaker babble)

**Comprehension run:**
- Speech-in-noise
- Clear speech
- Noise-only rest
- Speech-in-noise
- Silent Rest
- Clear speech

**Perception run:**
- Clear speech
- Silent Rest
- Speech-in-noise
- Clear speech
- Noise-only rest
- Speech-in-noise

Block ~14 s

Stimulus ~1.7 s

Go trials = false statements

Go trials = repeated sentence

**Signals:**
- Clear speech
- Speech embedded in noise
- Noise baseline
- Classic baseline
Behavioral data on Go trials

Brain-imaging data on Nogo trials

Global network
Speech processing > Silent baseline

• Task effect
  - Increased VWFA activation in comprehension compared to perception task
  - Small but significant VWFA activation in the perceptual task

• Noise effect
  - Adding noise into the signal made the task more difficult but did not lead to an increase of VWFA activity. The opposite tendency was found. Processing difficulty plays a role but not in the expected way → Disengagement of VWFA activity when speech is degraded?
    → Stronger propagation of brain activity to “non-phonological” systems when speech is clearly perceived?
• Non language auditory input induces “cross-model sensory suppression” within the visual system
• Using noise baseline artificially increases the effect size → can change the conclusion of the study (e.g., perception in noise)!
Anatomical correspondence between the voxels activated by spoken sentences and those activated by written words

Same or adjacent voxels?

Cohen et al., 2004
Anatomical correspondence between the voxels activated by spoken sentences and those activated by written words

**Localizer experiment**
- Aim: Localize the voxels that **respond to written words** in each individual subject
- Additional fMRI run: Visual presentation of **Words** or **Consonant strings**.
- Task: detection of a target stimulus “####” that appeared randomly
Visual word processing
Words > Consonant strings

L-Postcentral/precentral gyrus
L-Pars opercularis

L-Middle temporal gyrus
L-Inferior temporal, fusiform gyrus

• voxelwise, p < 0.001, uncorr.
• 50 most activated voxels (unthresholded)

Subject specific voxels within the search volume

Search Volume

z = -17

Subject-specific visual-vOT activation in speech processing
(> silent rest baseline)

Visual vOT = auditory vOT?
The degree of overlap between “visual vOT” and “auditory vOT”

Visual word processing
Words > Consonant strings

All speech > silent

Search volume

50 most activated voxels (unthresholded)

red: Visual vOT  green: Auditory vOT  yellow: Overlap

S20: 72% overlap

x = -50  y = -52  z = -16

Group average = 52% overlap
Underlying mechanisms of VWFA involvement in speech processing

Pattamadilok, C., Planton, S., & Bonnard, M., NeuroImage, 2019
Three hypotheses

**Orthographic tuning hypothesis** (Cohen & Dehaene 2005; Dehaene & Cohen, 2011).

The VWFA only contains neurons that encode orthographic inputs. Spoken input is first converted to an orthographic representation that activates these *written language-encoding neurons* (red dot) in a top-down fashion.
Orthographic tuning hypothesis

**Multimodal neurons hypothesis** (Price & Devlin, 2003; 2011)

The left-vOT contains **multimodal neurons** (blue dot) that encode language input regardless of its modality.
Orthographic tuning hypothesis

Multimodal neurons hypothesis

Heterogeneous neuronal populations hypothesis (Price & Devlin, 2003)

The left-vOT contains different populations of neurons that selectively encode written (red dot) and spoken (green dot) language inputs.
Properties of the neurons within the VWFA revealed by TMS-adaptation protocol

Main application: Causal role
If a brain area plays a causal role in the task under investigation, using TMS to interfere with the cortical excitability of the area should also modify task performance.

The effect of TMS can be inhibitory or facilitatory depending on stimulation parameters (e.g., frequency, intensity, timing, ...) and the initial state of the stimulated area (e.g., resting state, active, task-demands, ...).
**State-dependent TMS effects** (Silvanto et al. 2008): TMS effects are determined by initial neural activation state.

- By manipulating the state of neurons before application of TMS, one can selectively target specific, even spatially overlapping neural populations within the affected region, if those populations exist.
**Experimental paradigm & Predictions**

**Manipulation of the modality of language input**

- **Adaptation**
  - Task

**Experimental conditions**

<table>
<thead>
<tr>
<th>Orthographic</th>
<th>Multimodal</th>
<th>Subpopulations</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Brain with Orthographic" /></td>
<td><img src="image" alt="Brain with Multimodal" /></td>
<td><img src="image" alt="Brain with Subpopulations" /></td>
</tr>
</tbody>
</table>

- **TMS facilitatory effect**

<table>
<thead>
<tr>
<th>Orthographic</th>
<th>Multimodal</th>
<th>Subpopulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

- **Predictions**

<table>
<thead>
<tr>
<th>Orthographic</th>
<th>Multimodal</th>
<th>Subpopulations</th>
</tr>
</thead>
<tbody>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

- **Manipulation**

- **TMS**

- **Orthographic**

- **Multimodal**

- **Subpopulations**
Experimental protocol

N = 17 (8 women, 20-27 yrs, right-handed, native French speakers)

LEFT-vOT (VWFA) (Word > Fixation) ∩ MNI [-44, -58, -15, Jobard et al., 2003]

RIGHT-vOT (Control site)

Test phase (lexical decision)
- Double-pulse TMS (10Hz)
- 100 msec. post-stimulus onset
- 50% of the trials

Adaptation phase (passive listening or viewing)

TMS effect on task performance (Trials with TMS > Trials w/o TMS)
TMS & Neuronavigation system

fMRI localizer task
Words > Fixation

VWFA area on participant’s anatomical scan

VWFA area targeted by TMS during the task

Infrared camera tracking the participant’s head

TMS coil

Neuronavigation system
Linear mixed-effects model (RT data)
Fixed factors: TMS * Adaptation modality * Task modality * Lexicality * ROI
Random factors: Subjects and items
TMS * Adaptation modality * Task modality * ROI: F(1,11657) = 4.3, p < .05
### Experimental paradigm & Predictions

**Manipulation of the modality of language input**

- **Adaptation**
- **Task**

#### Experimental conditions

<table>
<thead>
<tr>
<th></th>
<th>word</th>
<th>word</th>
<th>word</th>
</tr>
</thead>
<tbody>
<tr>
<td>![Image 1](Image 1.png)</td>
<td>![Image 2](Image 2.png)</td>
<td>![Image 3](Image 3.png)</td>
<td>![Image 4](Image 4.png)</td>
</tr>
<tr>
<td>![Image 5](Image 5.png)</td>
<td>![Image 6](Image 6.png)</td>
<td>![Image 7](Image 7.png)</td>
<td>![Image 8](Image 8.png)</td>
</tr>
<tr>
<td>![Image 9](Image 9.png)</td>
<td>![Image 10](Image 10.png)</td>
<td>![Image 11](Image 11.png)</td>
<td>![Image 12](Image 12.png)</td>
</tr>
<tr>
<td>![Image 13](Image 13.png)</td>
<td>![Image 14](Image 14.png)</td>
<td>![Image 15](Image 15.png)</td>
<td>![Image 16](Image 16.png)</td>
</tr>
</tbody>
</table>
How does the VWFA respond to speech?

Heterogeneous neuronal populations hypothesis: functional segregated populations of neurons in the ventral visual system: written-language coding and spoken language-coding neuronal populations

Does not rule out the possibility that VWFA responses to spoken input could also be explained by a top-down activation of orthographic representations.
Beyond simple connections between the auditory and visual systems

Profound changes within the auditory and visual systems

Emergence of spoken language coding neurons in the visual pathway: Cortical reorganization induced by intensive training
- Task specific
- Depend on reading expertise (not in dyslexic or illiterates)

Functions =?
Fast communication between phonology and orthography in skilled readers and in high-level tasks?
Special thanks to...

Jean-Luc Anton
Bruno Nazarian
Julien Sein
Valérie Chanoine
Mireille Bonnard
Samuel Planton
Christophe Pallier