Morphological Processing in a Minimally Inflected Language

Yun-Hsuan Huang[1,2], Elisabeth Fonteneau[1,2], Caroline M. Whiting[1,2], Qing Cai[3,4] & William D. Marslen-Wilson[1,2]

[1]Department of Psychology, Cambridge, UK; [2] MRC Cognition and Brain Sciences Unit, Cambridge, UK
[3]East China Normal University, China; [4] INSERM U992 Cognitive Neuroimaing Unit, France

Introduction

It is widely assumed that combinatorial processes (morphological and syntactic processing) are the core computational processes of human language.

Cross-linguistic neuroimaging evidence suggests that spoken word comprehension engages two distinct but overlapping networks: combinatorial processes such as morphological inflections engage a left-lateralized fronto-temporal network in contrast to non-inflected words, which are supported by a more widely distributed bilateral fronto-temporal network. This evidence comes from richly inflected or moderately inflected languages, for example English [1], Arabic [2], and Polish [3]. We aimed to test for the presence of the dual network in a minimally inflected language, Mandarin Chinese.

We tested three types of disyllabic words varying in their combinatorial properties: compound words, derivation-like words (words whose second morpheme is a word forming suffix), and inflection-like verbs (verbs whose second morpheme is an aspectual marker). The dominant word type in Mandarin Chinese is compound words and the so-called derived and inflected words are much less common.

It is possible that similar patterns of activation to English will be observed in Mandarin Chinese – non-inflected words, such as compound and derived words, will engage the bilateral fronto-temporal network, and aspectual verbs will be supported by the left-lateralized fronto-temporal network, because they involve combinatorial processing. Alternatively, given that in other studies [4] derived words have been found to show decomposition in the left hemisphere, it is also possible that both derived and inflected words will engage a left-lateralized network in Mandarin Chinese.

Methods

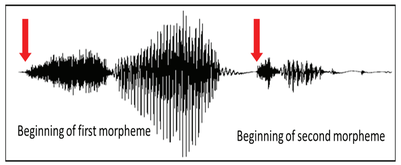
Stimuli - Three main types of disyllabic words, differing in the identity of the second morpheme were contrasted. 29 items per condition; 1 presentation for each block (4 presentations in total).

WORD TYPE	CATEGORY	2ND SYLLABLE	EXAMPLE	GLOSS
Compound	Verb	root morpheme	*hu* 'to exhale' *xi* 'to suck in'	to breath
	Noun	root morpheme	*gu* 'old' *shi* 'event'	story
Derived	Verb	verbalizing morpheme *hua*	*jing* 'pure' *hua*	to purify
	Noun	nominalizing morpheme *du*	*shi* 'wet' *du*	wetness
Inflected	Perfective	perfective marker *le*	*chang* 'sing' *le*	have sung
	Continuous	continuous marker *zhe*	*mo* 'touch' *Zhe*	touching

Task - 22 native speakers of Mandarin Chinese passively listened to the stimuli and on 10% of the trials performed a one-back semantic task.

Acquisition - Simultaneous EEG-MEG data (306-channel Vectorview system and a 70-channel EEG cap).
Two sets of epochs were generated: from word onset to 900msec after the word onset; and 100 msec before and 600 msec after the onset of the second morpheme.

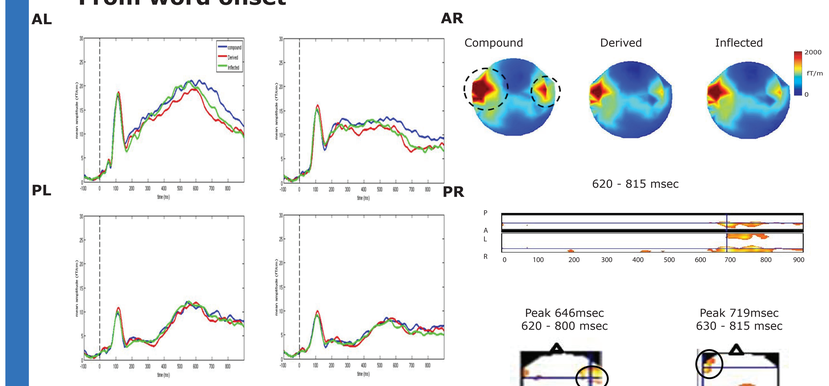
Source Reconstruction - Three-layer boundary element mdoel (Freesurfer) using participants' MRI scans. L2 Minimum Norm Estimate [5] was used to compute the source.



Analyses – At each data point, sensor level: mass univariate (F-tests, 3 conditions) with SPM (Random Field Theory) on gradiometers (voxel threshold $p < .01$, cluster extent $p < .05$); source level: paired-sample t-tests with cluster-mass permutations (N = 10000, one-tailed, $p < .05$) [6] at each predefined ROI.

Sensor-level Results

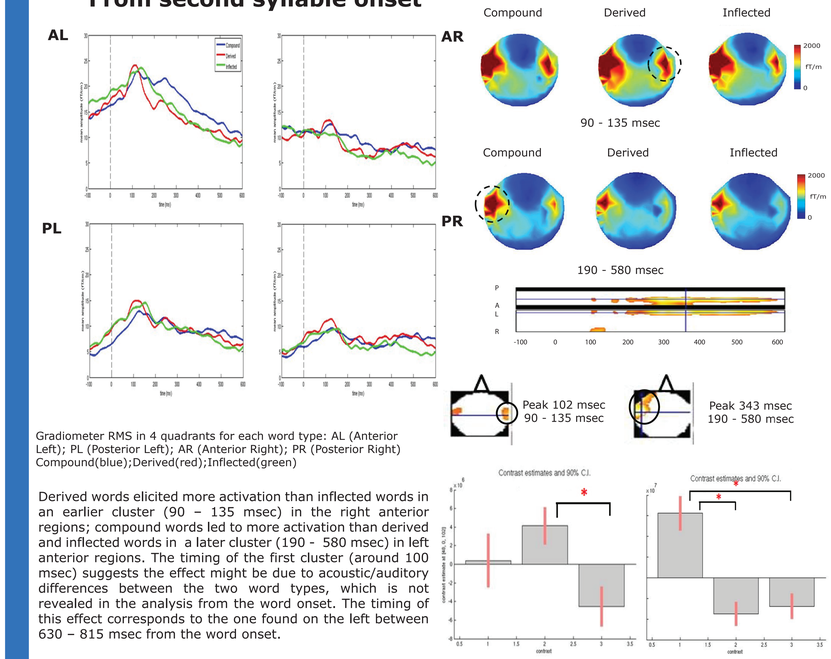
From word onset



Gradiometer RMS in 4 quadrants for each word type: AL(Anterior Left);PL (Posterior Left); AR(Anterior Right); PR(Posterior Right)
Compound(blue);Derived(red);Inflected(green)

620 – 800 msec after the onset of the word, in the right anterior regions, compound words elicited a larger amplitude of activity than derived and inflected words. Compound words also produced more activity than the other word types in 630 – 815 msec in the left anterior region. The late timing of the effect showed that there is no difference due to the first morpheme and that the effect was driven by the second morpheme, which disambiguates word type.

From second syllable onset

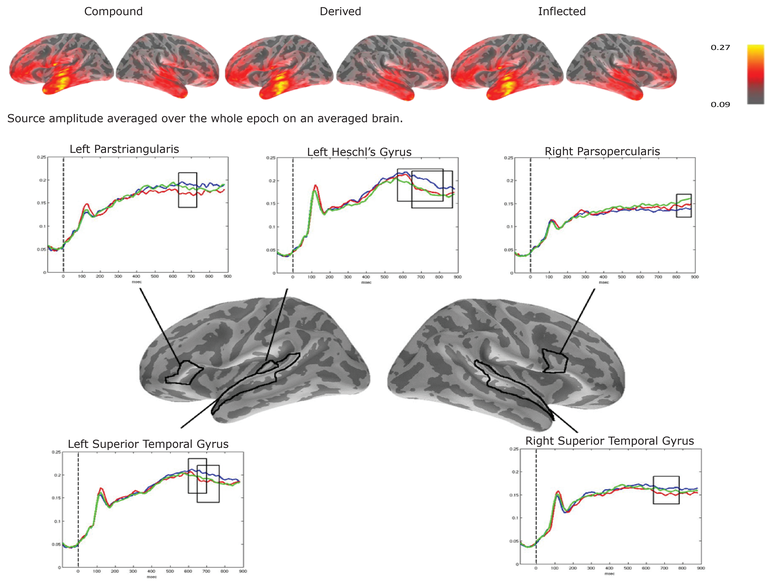


Gradiometer RMS in 4 quadrants for each word type: AL (Anterior Left); PL (Posterior Left); AR (Anterior Right); PR (Posterior Right)
Compound(blue);Derived(red);Inflected(green)

Derived words elicited more activation than inflected words in an earlier cluster (90 – 135 msec) in the right anterior regions; compound words led to more activation than derived and inflected words in a later cluster (190 - 580 msec) in left anterior regions. The timing of the first cluster (around 100 msec) suggests the effect might be due to acoustic/auditory differences between the two word types, which is not revealed in the analysis from the word onset. The timing of this effect corresponds to the one found on the left between 630 – 815 msec from the word onset.

Source-level Results

From word onset



Source amplitude averaged over the whole epoch on an averaged brain.

Timecourse plots for ROIs where significant or marginally significant clusters were found (rectangular windows).
Compound(blue); Derived(red); Inflected(green).

In the left hemisphere, compound words elicited more activation than derived and inflected words in Heschl's Gyrus and superior temporal regions; they also produced more activation than derived words in pars triangularis. All these differences were found late (around 620 – 780 msec). In the right hemisphere, inflected words produced more amplitude than compound words in pars opercularis (800 – 880 msec) and in superior temporal gyrus compound words produced more amplitude of activation than derived words. The timing of these differences are observed late in the second half of the epoch. This suggests the differences are due to the differences in the second morpheme, the distinguishing point of the three word types. (Apart from the two clusters in left Heschl's Gyrus, $p < .05$, all other clusters reported are marginally significant, $.05 < p < .10$)

Conclusions

In these preliminary analyses, sensor and source-level results from word onset and the sensor-level results from second syllable onset together showed that compounds produced more activity in left and right frontal and temporal areas than derived and inflected words in Mandarin Chinese around 200 msec after second syllable onset. This effect may reflect differences in compositionality between the word types: for derived and inflected forms the meaning of the word can be accessed by decomposing the word; whereas the meaning of a compound word cannot be obtained from the meaning of the constituent morphemes but requires retrieval of the whole word.

In addition, sensor-level results from second syllable onset revealed that around 100 msec after onset, derived words produced more activity in right anterior regions than inflected words. This effect might be due to the tonal difference in the second morpheme between the conditions: it is a falling lexical tone for derived words and a non-lexical tone for inflected words.

Spoken word processing in Mandarin Chinese mainly engages the left fronto-temporal network, although compound words show some engagement of the bilateral fronto-temporal network relative to derived and inflected words. Further research is required to determine how far these patterns parallel the results seen in languages like English.

References

1. Bozic, M., Tyler, L. K., Ives, D. T., Randall, B., & Marslen-Wilson, W. D. (2010). Bihemispheric foundations for human speech comprehension. *Proceedings of the National Academy of Sciences, 107(40),* 17439-17444.
2. Boudelaa, S., Pulvermüller, F., Hauk, O., Shtyrov, Y., & Marslen-Wilson, W. (2010). Arabic morphology in the neural language system. *Journal of Cognitive Neuroscience, 22(5),* 998-1010.
3. Szlachta, Z., Bozic, M., Jelowicka, A., & Marslen-Wilson, W. D. (2012). Neurocognitive dimensions of lexical complexity in Polish. *Brain and Language, 121(3),* 219-225.
4. Solomyak, O. & Marantz, A. (2009) Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience, 22:9,* 2042 - 2057.
5. Hämäläinen, M. S. & Ilmoniemi, R. J. (1994). Interpreting magnetic fields of the brain: minimum norm estimates. *Medical and Biomedical Engineering and Computing, 32,* 35 – 42.
6. Maris, E. Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods, 164,* 177-190.

yhh25@cam.ac.uk