Supporting Appendix

PNAS MS 2010-00531R: " Bi-hemispheric foundations for human speech comprehension"

Contents:

1. Musical Rain

Musical Rain – method of production.

The production of musical rain (17) uses a technique similar to that used to generate synthetic vowel sounds, by summing together continuous streams of four damped sinusoids. Typically the frequency of the sinusoids represents the formant frequencies and the repetition rate represents the pitch of the vowel. However, for the production of MuR, the carrier frequencies and the repetition rates of each of these sinusoids are randomized: over about a one octave range for the carrier frequency, and over a 20-ms period for the repetition rate. Temporal envelopes are extracted by taking the absolute value of the Hilbert transform for each speech token, and smoothed by low-pass filtering at 20 Hz. The temporal envelopes are then used to modulate the envelopes of the corresponding MuR tokens. The RMS level of the MuR is also matched to that of the speech, producing stimuli in which the long-term spectro-temporal distribution of energy is matched to that of the corresponding speech stimuli (Figure S1).

Figure S1

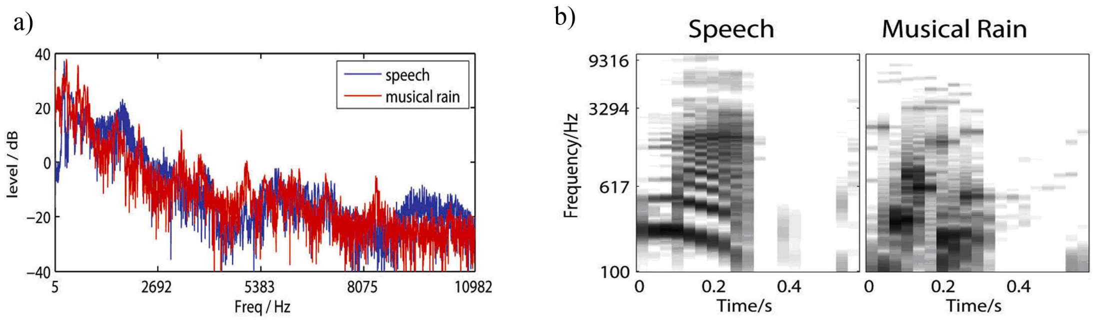

Fig S1: Properties of Musical Rain (MuR tokens and matched speech samples

(a) Fourier transforms of a speech token and the corresponding Musical Rain (MuR) sample. The similarity of the two Fourier transforms shows that the frequency spectra are well matched. (b) Spectrograms of the same tokens of speech and MuR, showing how modulating the envelope of the MuR produces similar patterns in the distribution of energy over the duration of the sound. The formants are clearly visible in the speech spectrogram but are absent in the MuR spectrogram.

2. MLM analyses

MLM analysis procedures

The MLM procedure used here is derived from the canonical variate analysis (CVA), a multivariate technique that incorporates a priori information provided by the GLM, and produces eigencomponents that minimise the within group variation while maximising the between group variation. It is performed on a multidimensional contrast (e.g., like that used to specify F-test effects of interest), comprising parameter estimates for the five experimental conditions as well as subject specific columns to remove the between subject variability. The resulting eigencomponents are orthogonal to each other, and show the greatest difference between the parameter estimates in a multivariate network. The method is purely descriptive, aimed to explore the relationship between the predictors and the data, and to help determine linear combinations of the predictors that best summarize the correlation structure.

Figure S2

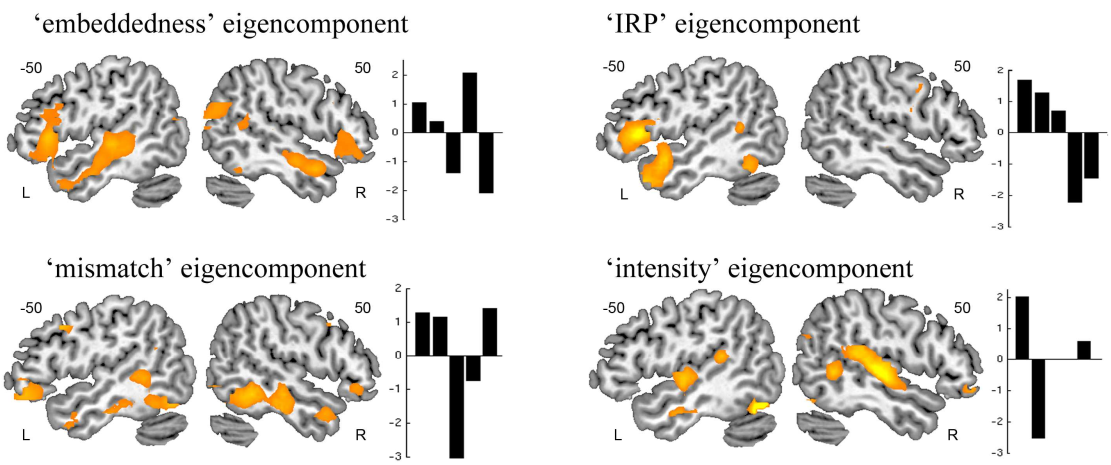

Fig S2. MLM Eigencomponents: Four eigencomponents were obtained in total. In addition to the two described in the manuscript ('embeddedness', which account for 59% of variance; and 'IRP', which accounts for 11%), another component dissociates words in the *blend* condition from the other four conditions, explaining 23% of variance. This is likely to reflect the activation triggered by 'mismatch': the presence of the IRP ending in *blend* triggers combinatorial processing, but the remaining *blen* is not a legitimate word in English and cannot be mapped onto stored lexical representation. The failure of this attempted mapping may trigger the strongly bilateral fronto-temporal activation associated with this component, similar to the one observed for the 'general processing demands/embeddedness' component.

The fourth component dissociates words from *prayed* and *claim* conditions from the words in the *trade* condition, and explains 8% of variance. Overlaid on the brain, activation is seen primarily in auditory cortices in both hemispheres. This may reflect the marginally significant variation in the acoustic intensity of these words sets, measured in dB [means per condition: *prayed* 59.36; *trade* 56.55; *blend* 57.54; *claim* 59.50; *dream* 59.74; $F(4,195)=2.08$, $p=.084$]. All components are shown thresholded at $t>1$.

Bar graphs show how the original predictors (conditions) weight on the linear combinations expressed in an eigencomponent. They are in the unit of the beta images, and can be compared across conditions within a given component. The order of bars in each graph is *prayed, trade, blend, claim, dream* (from left to right).

3. Univariate analyses: effects at lower threshold

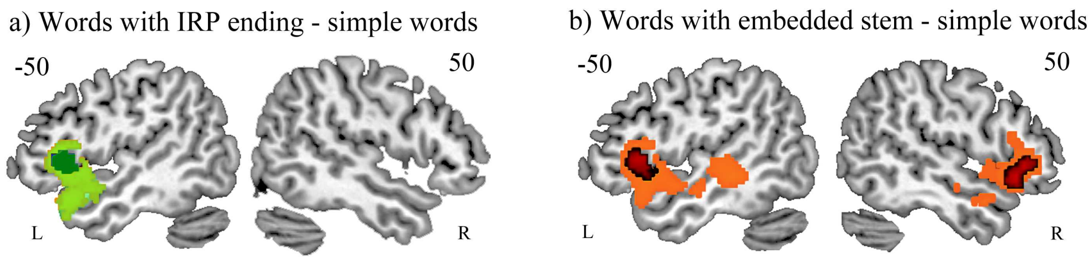

Fig S3. Univariate analyses: Activation at p<.001 and p<.01 voxel-level thresholds, significant at p<.05 cluster-level corrected for multiple comparisons, for (a) words with IRP endings minus simple words and (b) words with embedded stem minus simple words. Lighter colours show the activation at the lower (p<.01) threshold.

4. Overlap in LIFG activation

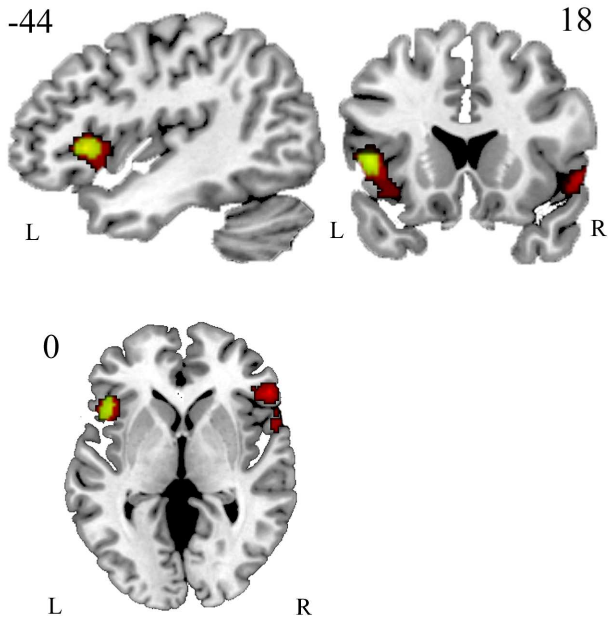

Fig S4. Activation overlap: Overlap between the IRP-related activation and the embedded stem–related activation. IRP-related activation (BA 45, light green/yellow) is fully included within the L embedded stem–related activation (BA 45/47, dark red).

5. Analyses of cohort effects and of non-initial embeddedness

a) Cohort properties of stimuli

In addition to the analyses conducted in terms of onset-embedded competitor stems, we checked the cohort properties of the stimuli, matching these across conditions and testing for effects in the behavioural and neuro-imaging results.

(i) Focusing on the relative weight of cohort competitors as estimated by their frequency, we used the CELEX database (Baayen, Piepenbrock, & Gulikers, 1995) to obtain a measure of the summed frequency of all the cohort members for each test items (cohort members were defined as words (of whatever length) sharing the initial CV/CCV of the word in question). An ANOVA showed that there was no difference in summed cohort frequencies across the five conditions [$F(4,195)<1$].

(ii) As a more specific index of cohort competition for single items, we looked at the cohort frequency ratio, defined as the ratio of frequency of the specific word to the summed cohort frequency for that word. Here again there was no difference across conditions [$F(4,195)=1.5$, $p>.1$].

(iii) In addition, we correlated the summed cohort frequency and the cohort ratio with the behavioural scores (reaction times) from the scanner. The results showed no significant correlation.

(iv) Finally, we entered the cohort ratio, on an item-by-item basis, as a parametric modulator in the imaging analysis. There were no regions that showed significant modulation by this variable.

b) Non-initial embeddedness

Non-initial embeddedness is a common occurrence in language, and this is reflected in the stimuli: several words in each condition contained mid or late (rhyme) embedded words. These were often frequent short words such as *end, if, in,, aim,* etc, which appear in words from all five conditions, often more than one (e.g. *lay, lame, aim* in *claim*). We tested whether the numbers and the frequency of these words affect the processing of the word they are embedded in.

(i) Using the CELEX database, we extracted all non-initial embedded words and their frequencies for each of the test items. We then correlated the behavioral scores (reaction times) from the scanner with the number of non-initial embedded words for each item, as well as with the ratio of the frequency of the specific word to the summed frequency of all its late embedded words (labeled the non-initial embeddedness ratio). There were no significant correlations.

(ii) We tested the distribution of the non-initial embeddedness ratio across the five conditions. An ANOVA showed no difference across conditions [$F(4,195)<1$].

(iii) We also entered the non-initial embeddedness ratio as a parametric modulator in the imaging analysis, on an item-by-item basis. No regions showed significant modulation by this variable.

Table S1

a) Regions supporting complex acoustic processing; b) regions showing hemispheric laterality effects. Throughout, results were thresholded at $p < .001$ voxel level and clusters that survived $p < .05$ corrected for multiple comparisons were considered significant. MNI coordinates are reported. The highest three peaks within a cluster are shown, with the most significant marked in bold. MuR: Musical Rain.

Regions	*Cluster*		*Voxel*	*MNI Coordinates*		
	$P_{correct}$	*Extent*	*Z*	*x*	*y*	*z*
(a) MuR - null						
L Superior temporal gyrus (BA 42)	0.000	1575	4.49	-66	-30	14
L Superior temporal gyrus (BA 22)			4.19	-44	-40	18
L Heschl's gyrus (BA 41)			4.10	-44	-28	16
R Superior temporal gyrus (BA 42)	0.000	1190	3.98	68	-34	14
R Superior temporal gyrus (BA 22)			3.95	62	-16	4
R Superior temporal gyrus (BA 22)			3.80	58	-36	12
(b) Laterality						
L Insula	0.000	819	4.69	-40	0	12
L Heschl's gyrus (BA 41)			4.56	-38	-24	10
L Heschl's gyrus (BA 41)			4.29	-38	-34	14

Table S2

a) Regions supporting speech-driven lexical processing; b) hemispheric laterality of this process.

Regions	*Cluster*		*Voxel*	*MNI Coordinates*		
	$P_{correct}$	*Extent*	*Z*	*x*	*y*	*z*
<u>**(a) *Words - MuR***</u>						
L Middle temporal gyrus (BA 21)	0.000	1710	5.41	-64	-14	-6
L Middle temporal gyrus (BA 21)			4.58	-66	-26	0
L Anterior temporal cortex (BA 38)			4.48	-44	16	-32
R Middle temporal gyrus (BA 21)	0.000	536	4.38	60	10	-16
R Middle temporal gyrus (BA 21)			4.26	64	2	-16
R Middle temporal gyrus (BA 21)			3.90	68	-10	-4
L Fusiform gyrus (BA 20)	0.000	477	4.65	-38	-22	-18
L Fusiform gyrus (BA 37)			4.47	-32	-36	-20
L Hippocampus			4.42	-40	-34	-4
L Angular gyrus (BA 39)	0.000	362	4.05	-42	-70	26
L Angular gyrus (BA 39)			3.95	-36	-64	26
L Middle occipital gyrus (BA 19)			3.66	-40	-80	32
R Hippocampus	0.000	264	5.11	34	-34	-2
R Hippocampus			4.48	38	-30	-8
R Fusiform gyrus (BA 37)			4.24	36	-40	-8
R Angular gyrus (BA 39)	0.002	209	3.85	54	-68	28
R Angular gyrus (BA 39)			3.82	52	-74	16

R Angular gyrus (BA 39)			3.51	48	-74	32
R Anterior cingulate (BA 32)	0.005	168	4.30	8	22	-10
L Anterior cingulate (BA 32)			4.22	-8	26	-10
R Anterior cingulate (BA 32)			3.66	12	30	-8
(b) Laterality						
L Precentral gyrus (BA 6)	0.000	127	4.33	-40	8	40
L Pars opercularis (BA 44)			4.02	-52	10	34
L angular gyrus (BA 39)	0.000	110	4.51	-40	-74	26
L Pars orbitalis (BA 47)	0.000	104	3.93	-42	28	-14
L Pars orbitalis (BA 47)			3.77	-50	24	-10
L Pars orbitalis (BA 47)			3.51	-36	30	-8
L Pars Triangularis (BA 45)	0.021	48	3.88	-54	24	4

Table S3

Regions showing processing differences as a function of type of lexical complexity: a) Simple words; b) Words with a potential inflectional ending (Inflectional Rhyme Pattern); c) Words with embedded stems. To show the full extent of the IRP-related and stem–related activation, in b) and c) we also report the subsequent most significant peaks from the lower threshold of p<.01. These peaks are marked by an asterisk.

Regions	*Cluster*		*Voxel*	*MNI Coordinates*		
	$P_{correct}$	*Extent*	*Z*	*x*	*y*	*z*
(a) Simple -MuR						
L Middle temporal gyrus (BA 21)	0.000	486	5.02	-60	-16	-6
L Middle temporal gyrus (BA 21)			4.29	-64	-6	-4
L Superior temporal gyrus (BA 22)			4.19	-66	-22	-2
R Medial temporal gyrus (BA 21)	0.009	163	4.92	38	-40	-8
R Parahippocampus (BA 37)			4.29	40	-32	-4
(b) IRP present - simple						
LIFC Pars triangularis (BA 45)	0.015	111	3.82	-48	18	4
L superior temporal pole (BA 38)*			2.94	-58	12	-10
L Superior temporal gyrus (BA22)*			2.85	-46	6	-22
(c) Embedded –simple						
RIFC Pars orbitalis (BA 47)	0.000	321	4.50	54	30	0
RIFC Pars orbitalis (BA 47)			4.19	54	22	-8
R superior temporal pole (BA 38)			3.85	56	10	-4
RIFC Pars orbitalis (BA 47) *			3.79	46	28	-4
R Superior temporal gyrus (BA22)*			3.53	56	-2	-10

R Middle temporal gyrus (BA 21)*			3.52	66	-18	-6
LIFC Pars orbitalis (BA 47)	0.000	247	4.43	-44	22	2
LIFC Pars orbitalis (BA 47)			3.31	-36	18	-10
L Middle temporal gyrus (BA 21)*			3.79	-48	-34	-6
L Superior temporal gyrus (BA22)*			3.56	-66	-40	6
L Middle temporal gyrus (BA 21)*			3.39	-60	-24	-4

Table S4

Regions showing modulation by lexical competition.

Regions	*Cluster*		*Voxel*	*MNI Coordinates*		
	$P_{correct}$	*Extent*	*Z*	*x*	*y*	*z*
LIFC Pars orbitalis (BA 47)	0.013	143	4.02	-40	14	-2
L Insula			3.19	-38	6	0
R Insula	0.065*	87	3.98	46	6	-4
RIFC Pars orbitalis (BA 47)			3.61	42	16	-8

* p=.05 at voxel level threshold of p<.005

10. Experimental stimuli for Conditions 1-5 (British English)

Cond.1	Cond.2	Cond.3	Cond.4	Cond.5
bashed	beard	bind	barge	brisk
blamed	brand	bird	binge	bulb
blared	card	bleed	claim	bulge
blurred	chest	blend	clasp	cage
blushed	cleft	blind	coin	chain
boiled	crest	bond	crone	clean
bowed	crude	breed	cusp	cream
cared	deft	bride	drive	cringe
chewed	fade	broad	film	crown
creaked	fast	cloud	frame	deem
cried	feed	code	gain	desk
dared	fend	creed	game	dream
fried	feud	crust	grain	dusk
glared	fold	deed	grange	fringe
gleamed	fund	fiend	grasp	gown
hurled	gold	fist	grove	green
jeered	grade	fraud	gunge	harm
joined	graft	frost	helm	hinge
kicked	grand	ghost	keen	husk
laid	guide	grind	kiln	learn
mashed	hide	guard	lawn	lounge
peered	jade	hard	mosque	lunge
prayed	mild	hound	nine	nerve
purred	mould	lard	palm	plunge
roared	nude	lend	phone	prove
rolled	proud	loud	prime	prune
sailed	raid	nest	rage	queen
saved	rend	pond	range	rasp
shared	ride	rind	rhyme	realm
showed	seed	shade	serve	salve
shunned	shrewd	shield	shelve	shame
sniffed	slide	shift	shine	shave
stared	tend	shroud	sign	sleeve
stayed	trade	swede	singe	task
stirred	tuft	third	spine	term
thawed	tweed	vest	sponge	vague
walked	wand	void	swarm	valve
warned	ward	wound	tinge	wasp

washed	weird	yard	tone	wisp
yawned	wind	yield	twinge	zone