

## Short Communication

# Cross-linguistic parallels in processing derivational morphology: Evidence from Polish



Mirjana Bozic, Zanna Szlachta, William D. Marslen-Wilson\*

Department of Psychology, University of Cambridge, UK  
MRC Cognition and Brain Sciences Unit, Cambridge, UK

## ARTICLE INFO

## Article history:

Accepted 1 September 2013  
Available online 14 October 2013

## Keywords:

Neurobiology of language  
Morphology  
Derivation  
Polish  
fMRI  
Cross-linguistic

## ABSTRACT

Neuroimaging evidence in English suggests that the neurocognitive processing of derivationally complex words primarily reflects their properties as whole forms. The current experiment provides a cross-linguistic examination of these proposals by investigating the processing of derivationally complex words in the rich morphological system of Polish. Within the framework of a dual language system approach, we asked whether there is evidence for decompositional processing of derivationally complex Polish stems – reflected in the activation of a linguistically specific decompositional system in the left hemisphere – or for increased competition between the derived stem and its embedded base stem in the bilateral system. The results showed activation in the bilateral system and no evidence for selective engagement of the left hemisphere decompositional system. This provides a cross-linguistic validation for the hypothesis that the neurocognitive processing of derived stems primarily reflects their properties as stored forms.

© 2013 The Authors. Published by Elsevier Inc. Open access under [CC BY-NC-ND license](http://creativecommons.org/licenses/by-nc-nd/4.0/).

## 1. Introduction

Derivational morphology is one of the major mechanisms of word formation cross-linguistically. Across many different languages, word stems (e.g. *happy*, *speak*) are combined with derivational morphemes (e.g. *-ness*, *-er* in English) to express new meanings and create new entries in the mental lexicon (e.g., *happiness*, *speaker*). On-line language comprehension requires the listener to recognise the

### 1.1. Background

Recent neuroimaging studies in English and Polish (Bozic et al., 2010; Szlachta, Bozic, Jelowicka, & Marslen-Wilson, 2012) provide evidence that spoken language comprehension engages two joint but functionally distinguishable neurobiological systems: a distributed bilateral system, which supports general perceptual and interpretative processes underpinning speech comprehension,

primarily reflects their properties as whole forms (Bozic, Tyler, Su, Wingfield, & Marslen-Wilson, 2013), and does not invoke the left-lateralised decompositional mechanisms that are central to the processing of inflectionally complex forms (Bozic, Tyler, Ives, Randall, & Marslen-Wilson, 2010; Marslen-Wilson & Tyler, 2007). Here we investigate the processing of derivationally complex words in Polish, to test the cross-linguistic validity of the claims made for derived forms in English.

larly inflected forms – English words like *played* or *yards* and Polish words like *sklepem* ‘shop, Instr’ or *czytam* ‘I read’. These words, combining a stem with an inflectional suffix, selectively activate the inferior frontal gyrus (IFG) on the left but not on the right.

In these earlier studies, we contrasted this linguistically-specific source of processing complexity with more general perceptual processing complexity generated by the presence of an onset-embedded stem – for example, English words like *claim*, with the embedded pseudostem *clay*, or Polish words like *kwitnie* ‘[it] blooms’ with the embedded pseudostem *kwit* ‘receipt’. The perceptual competition generated by these forms, as competing members of the same word-initial cohort (Marslen-Wilson, 1987; Zhuang, Tyler, Randall, Stamatakis, & Marslen-Wilson, 2012), triggered a symmetric bilateral fronto-temporal pattern of activation, consistent with the engagement of the hypothesized bihemispheric system.

\* Corresponding author. Address: Neurolex Group, Department of Psychology, Downing Site, University of Cambridge, Cambridge CB2 3EB, UK. Fax: +44 (0) 1223 333564.

E-mail address: [wdm10@cam.ac.uk](mailto:wdm10@cam.ac.uk) (W.D. Marslen-Wilson).

In a subsequent study (Bozic et al., 2013) we addressed the question of whether derivationally complex words in English engage the LH system in the same way as inflected forms, which would imply similarly decompositional processing and storage, or whether (and to what extent) they engage the distributed bihemispheric system, consistent with whole-form, non-compositional accounts. These questions relate to long-standing issues in the psycholinguistic literature about the role of morphemic structure in the processing and representation of derivationally complex words, with some accounts arguing for across-the-board stem-suffix decomposition (e.g., Taft, 2004); others emphasising the role of semantic compositionality and affix productivity in determining whether a complex word is stored in decompositional format (Bertram, Schreuder, & Baayen, 2000; Marslen-Wilson, Tyler, Waksler, & Older, 1994) or as whole forms (e.g., Butterworth, 1983) but with preserved morphological structure for transparent words with productive suffixes (Clahsen, Sonnenstuhl, & Blevins, 2003; Marslen-Wilson, 2007).

To address these issues in a neurocognitive context, using fMRI, the Bozic et al. (2013) study systematically varied the semantic transparency and the affix productivity of sets of English derived words, forming a gradient of potential decompositionality from semantically opaque forms with either non-productive or productive suffixes, such as *breadth* or *archer*, to semantically transparent forms with a productive suffix, such as *bravely* or *farmer*, using morphologically simple words like *giraffe* as controls. The results showed no evidence for selective activation of the LH system that supports combinatorial processing, even for the potentially most decomposable *bravely* or *farmer* forms. This contrasted strongly with the distinctive left-lateralised decompositional processes seen for English regular inflection. Instead, we saw increased activity in the distributed bilateral system, which was strongest for non-compositional derived words such as *breadth* or *archer*. This activation was primarily driven by the properties of the derived form as a whole and reflected the demands associated with perceptual competition between the derived words and their onset-embedded stem or pseudostem.

It was noteworthy, however, that no increase in activation was seen for the *bravely*-type transparent words, which patterned with simple forms like *giraffe*. This suggests that transparently derived words with productive affixes are not competitors to their embedded stems, implying a structured overlap between the two lexical representations (such that evidence for *brave* is also treated as potential evidence for *bravely*). This is consistent with the view that associates derivational word-formation with stored whole word representations, but where constituent morphological structure may be encoded for semantically transparent forms with productive affixes (Bozic & Marslen-Wilson, 2010; Clahsen et al., 2003; Marslen-Wilson, 2007).

### 1.2. Cross-linguistic implications

The results from English suggest that the processing of derived forms is driven by the demands of accessing a whole word representation in the presence of competing alternatives, and is primarily supported by the distributed bilateral system. This points to a strong distinction between inflectional and derivational morphology, with inflectional morphemes triggering decompositional linguistic processes, while derivational morphemes play their role as part of the whole word representations for each derived lexeme.

The cross-linguistic applicability of such a claim needs to be carefully examined, given the properties of inflectional and derivational morphological systems in English. In particular, inflectional morphology in English has a limited distribution, especially in the nominal system where the only inflectional morpheme is the plural {-s}. This means that the derived stem (base stem plus

derivational affix) and the whole word are generally the same phonological form, with no additional linguistic complexities. This contrasts with languages like Polish, where every noun, whether derivationally complex or not, is inflectionally marked for case, number, and gender. The neuter nominal form *badanie*, 'inspection', for example, from *badac*, 'to inspect', breaks down into the derived stem *bad-ani*, made up of the verbal root {*bad-*} with the derivational suffix {-*ani-*}, combined with the inflectional morpheme {-*e*} that marks the nominative case for a neuter singular noun.

On the view that inflectional morphemes do engage decompositional LH systems, and given that all derivational forms in Polish, whether verbs or nouns, carry inflectional markers, this means that derived stems in Polish will occur in a much more decompositional and left-lateralised processing environment than in English. In this study we ask whether, nonetheless, derivational complexity in Polish has the same neurocognitive hallmarks as in English, with no selective involvement of LH systems when derivational morphemes are present, and where the chief processing correlates reflect perceptual competition between the derived form and its onset-embedded base stem (or pseudo-stem),<sup>1</sup> engaging bihemispheric processing systems.

### 1.3. Derivational complexity in Polish: Experimental considerations

The rich derivational system in Polish contains many nouns, typically formed using derivational suffixes. Here we focus on a set of four regular and productive suffixes ({-*k*}, {-*anie*}/{-*enie*}, {-*ość*} and {-*arz*}), on the assumption that these are most likely to show decompositional processing similar to inflected words.

The suffix {-*k*} is a diminutive suffix that makes a noun-to-noun modification, and indicates small size or positive emotional value. It is highly productive and can be applied to almost any noun, with the form of the suffix varying according to the gender of the base word – e.g., [-*ek*] or [-*ik*] for masculine nouns, as in *papierek*, 'little piece of paper', from *papier*, 'piece of paper'; [-*ka*] for feminine nouns, as in *żabka*, 'little frog', from *żaba*, 'frog'; [-*ko*] for neuter nouns, as in *pudełko*, 'little box', from *pułto*, 'box'. There are occasional (regular) stem alternations during this modification (as in *pułto-pudełko*). The suffix {-*anie*}/{-*enie*} transforms verbs into nouns, similar to the English gerund (e.g., {-*ing*} in *swimming*), and is modulated by the phonological properties of the verb stem, as in *czytanie*, 'reading', from *czytać*, 'to read'; or *karmienie*, 'feeding', from *karmić*, 'to feed'. It applies to almost all verbs, apart from modals, and the resulting derived forms are usually transparent in meaning. The suffix {-*ość*} transforms adjectives into nouns, similar to English *-ness* (as in *happiness*). It is also a common suffix, and applies to almost all adjectives – for example, *młodość*, 'youth', from *młody*, 'young', and *życzliwość*, 'kindness', from *życzliwy*, 'kind'. The final suffix, {-*arz*}, can attach to verbs or nouns, with effects similar to the English agentive suffix {-*er*}, as in *worker*. An example of deverbal derivation is *malarz*, 'painter', from *malować*, 'to paint'. Denominal derivations include forms such as *lalkarz*, 'puppet maker', from *lalka*, 'puppet'.<sup>2</sup>

We used these four suffixes to create Polish versions of three critical conditions from the earlier study in English (Bozic et al., 2013). These were a semantically transparent productive affix condition (analogous to the *bravely/farmer* set in English), a semantically opaque productive affix condition (analogous to the opaque

<sup>1</sup> Each mapping onto its own whole-word representation

<sup>2</sup> Both {-*ość*} and {-*arz*} forms have a null morpheme in the nominative case but acquire overt inflectional morphemes in other cases. As shown in Szlachta et al. (2012), Polish nouns engage LH fronto-temporal systems linked to decompositional processing, whether or not the case inflection is overtly marked. We therefore expect all stimuli here to be treated as equivalently inflectionally complex.

**Table 1**  
Experimental conditions and sample stimuli (embedded stems, where present, are in parentheses).

Condition	N	Suffix	Example
Semantically transparent	60	-k- -anie/-enie -ość -arz	<i>pudełko</i> , 'little box' ( <i>pu</i> łto, 'box') <i>czytanie</i> , 'reading' ( <i>czyta</i> ć, 'to read') <i>młodość</i> , 'youth', <i>młody</i> , 'young' <i>lalkarz</i> , 'puppet maker', <i>lalka</i> , 'puppet'
Semantically opaque	60	-k- -anie/-enie -ość -arz	<i>kanapka</i> , 'sandwich' ( <i>kanapa</i> , 'couch') <i>kazanie</i> , 'sermon' ( <i>kaza</i> ć, 'tell') <i>mdłość</i> , 'nausea' ( <i>md</i> ły, 'fuzzy') <i>sekretarz</i> , 'secretary' ( <i>sekre</i> t, 'secret')
Simple	60	No suffix	<i>kapusta</i> , 'cabbage' <i>chirurg</i> , 'surgeon' <i>gitara</i> , 'guitar' <i>telewizor</i> , 'TV'

productive *archer* set in English), and a derivationally simple condition, where the stimuli contain neither a potential stem nor a derivational suffix (analogous to the English *giraffe* set) (Table 1).<sup>3</sup> An example of the semantically opaque condition, where the meaning of the derived word is unrelated to the meaning of the stem, is the form *smoczek*, 'dummy/pacifier', where the pseudostem *smok* means 'dragon' and the phonological change *k->cz* is a regular diminutive alternation. An example of the simple condition is the word *kapusta*, 'cabbage', containing neither an embedded stem nor a derivational morpheme (although bearing an inflectional suffix, the nominative feminine {-a}).

This set of contrasts allows us to test two hypotheses about the processing and representation of Polish derived stems. If the markedly more decompositional environment of Polish – due to the prevalence and richness of Polish inflectional morphology – affects the processing of derived stems such that they are also processed and represented decompositionally, then semantically transparent (and possibly also opaque) derivations should elicit increased activation in the LH fronto-temporal system compared to non-derived words. This result would differ from the English findings and point to significant differences between language processing mechanisms in different linguistic environments. On the other hand, if the results reflect demands on general perceptual processing due to increased cohort-type competition between a derived stem and its onset-embedded base stem (or pseudostem), similar to English derivations, then we should see bilateral effects for opaque and transparent derived words, but not accompanied by differential left-lateralized fronto-temporal effects.

These hypotheses were examined in an fMRI study, run on 22 native speakers of Polish, and using sparse imaging to avoid scanner noise during stimulus presentation. Participants were asked to listen attentively to the stimuli and to perform an occasional one-back memory task, intended to maintain attention. To isolate lexical processing from lower-level auditory processing, we used musical rain (MuR) – an acoustic baseline closely matched to speech in terms of its temporal envelope and overall energy, but which does not induce a phonetic percept.

## 2. Results

Based on previous research (Binder et al., 1997; Bozic et al., 2010; Szlachta et al., 2012; Tyler & Marslen-Wilson, 2008) we focused on bilateral fronto-temporal areas as the volume of interest for the analyses. A mask consisting of bilateral temporal lobes

(superior, middle and inferior temporal gyri, angular gyrus), inferior frontal gyri (pars orbitalis, pars opercularis, pars triangularis), insula and the anterior cingulate was constructed using WFU Pickatlas.

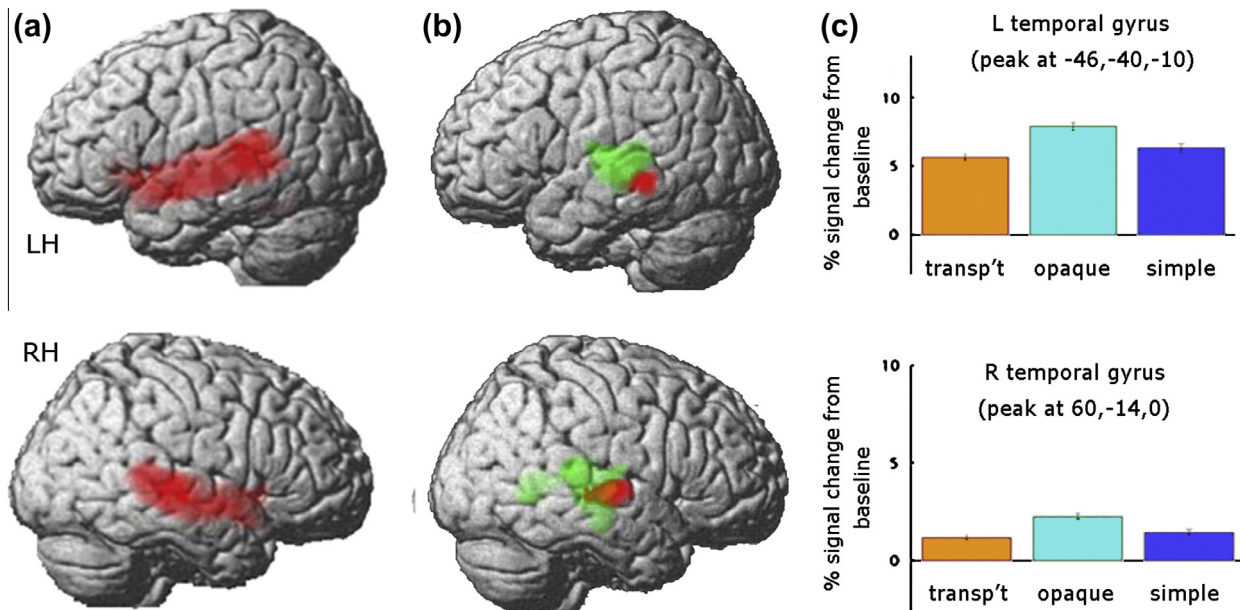
To establish the network that supports complex acoustic processing we subtracted null events from the acoustic baseline (MuR) and saw two clusters, one covering left middle and superior temporal gyri, Rolandic operculum and Heschl's gyrus, and the other the same areas on the right. Activation coordinates are given in the Supplement (Table S2).

The activity specifically related to lexical processing was extracted by contrasting words with the MuR baseline. This contrast revealed three significant clusters (Fig. 1a and Table 2a). The largest cluster included left middle and superior temporal gyri, left superior temporal pole and the LIFG (BA 47). The second cluster was in the right hemisphere, including middle and superior temporal gyri and superior temporal pole. The third cluster was in the left fusiform gyrus. The selective activation in LIFG (but not in RIFG) parallels the results in the earlier study (Szlachta et al., 2012), where all Polish forms (inflected verbs and nouns) activated left BA 47 (extending into BA 45). As in the previous study, all three word conditions activated the LIFG equally.

To test for differences between the three word conditions, the MuR baseline was subtracted from each condition and the resulting activations compared in a repeated measures ANOVA with added subject-specific effects. This analysis showed two significant clusters of differential activations (Fig. 1b (red overlay) and Table 2b). One was in the right middle and superior temporal gyri, the other in the left middle temporal gyrus. Plots of activation for the three conditions at the peak coordinates (Fig. 1c) show a comparable pattern in the two clusters – opaque words elicited stronger activation than either transparent or simple words. These results were confirmed by individual contrasts for each pair of conditions. No differential effects were seen in the LIFG, even at the lowest threshold.

To investigate whether these effects were linked to perceptual competition between the derived stems and their onset-embedded base stems, as suggested by previous findings in English, we tested for modulation of activity as a function of the degree of lexical competition. Competition was expressed as the ratio of the form frequencies of the derived word and its embedded stem or pseudostem, in transparent and opaque nouns (e.g., *żabka/żaba*; *kanapka/kanapa*), and entered as a parametric modulator. We found significant effects in bilateral middle and superior temporal gyri (Table 2c), in regions that overlapped with the sites of differential activation between the three conditions (see green overlays in Fig. 1b). Analysing the effects of competition in transparent and opaque conditions separately, however, we found that the opaque derivations drive the observed bilateral temporal activation, and

<sup>3</sup> Because of the absence of full distributional information for Polish morphemes, it was not feasible to also construct well-defined stimulus sets with unproductive affixes.



**Fig. 1.** (a) Lexical processing (all words – MuR baseline); (b) red: differential activation between the three conditions (transparent, opaque, and simple); green: modulation of activity as a function of lexical competition; and (c) condition signal plots at the peak coordinates of each (red) cluster ([60, –14, 0] and [–46, –40, –10]). All results are thresholded at  $p < .05$  cluster level corrected for multiple comparisons (red =  $p < .001$  voxel level; green =  $p < .01$  voxel level). Activations are shown rendered on the surface of a canonical brain.

**Table 2**

(a) Lexical processing; (b) significant differential activations between the three conditions; and (c) modulation of activity as a function of lexical competition.

Regions	Cluster		Z score	Peak voxel		
	p Corr.	Extent		x	y	z
<i>(a) Words – MR</i>						
<b>L middle temporal gyrus (BA 21)</b>	<b>0.000</b>	<b>4889</b>	<b>6.32</b>	<b>–62</b>	<b>–36</b>	<b>2</b>
L superior temporal pole (BA 38)			5.83	–56	8	–12
L middle temporal gyrus (BA 21)			5.67	–64	–24	–4
L superior temporal gyrus (BA 22)			5.61	–62	–14	–6
L inferior frontal gyrus (BA 47)			4.90	–34	26	–8
<b>R superior temporal gyrus (BA 22)</b>	<b>0.025</b>	<b>3697</b>	<b>5.99</b>	<b>66</b>	<b>–20</b>	<b>–2</b>
R middle temporal gyrus (BA 21)			5.51	62	–4	–8
R superior temporal pole (BA 38)			4.88	50	14	–10
<b>L Fusiform Gyrus (BA 37)</b>	<b>0.025</b>	<b>203</b>	<b>3.78</b>	<b>–36</b>	<b>–58</b>	<b>–18</b>
L Fusiform Gyrus (BA 37)			3.78	–36	–48	–22
L Fusiform Gyrus (BA 37)			3.57	–40	–58	–22
<i>(b) ANOVA</i>						
<b>R superior temporal gyrus (BA 22)</b>	<b>0.006</b>	<b>295</b>	<b>4.34</b>	<b>60</b>	<b>–14</b>	<b>0</b>
R STG/MTG (BA 22/21)			3.39	64	–26	–2
R Medial MTG (BA 21/48)			3.27	48	–20	–6
<b>L middle temporal gyrus (BA 21)</b>	<b>0.007</b>	<b>276</b>	<b>4.03</b>	<b>–46</b>	<b>–40</b>	<b>–10</b>
L R middle temporal gyrus (BA 21)			3.95	–50	–38	–8
L R middle temporal gyrus (BA 21)			3.81	–52	–36	–6
<i>(c) Lexical competition</i>						
<b>L middle temporal gyrus (BA 21)</b>	<b>0.012</b>	<b>1015</b>	<b>3.62</b>	<b>–56</b>	<b>–32</b>	<b>4</b>
L middle temporal gyrus (BA 21)			3.61	–52	–26	–4
L superior temporal gyrus (BA 22)			3.32	–50	–36	8
<b>R middle temporal gyrus (BA 21)</b>	<b>0.002</b>	<b>1444</b>	<b>3.57</b>	<b>40</b>	<b>–56</b>	<b>2</b>
R superior temporal gyrus (BA 22)				60	–12	14
R superior temporal gyrus (BA 22)				64	–20	–2

All results thresholded at  $p < .05$  cluster level corrected for multiple comparisons; (a), (b) voxel level  $p < .001$ ; (c) voxel level  $p < .01$ .

fully account for the competition effects reported in Table 2c and Fig. 1b. No significant competition effects were seen for the transparent forms, despite the fact that both conditions exhibited a similar range of competition values (.06–.96 for the transparent; .02–.99 for the opaque), with no difference between them [ $F(1, 119) = .387, p > .1$ ].

### 3. Discussion

This experiment investigated the processing consequences of derivational complexity in Polish, providing a cross-linguistic perspective on the neurocognitive processes observed for English derivation. We compared the processing of derivationally complex



transparent and opaque words with non-derived words, in a context where all forms are inflectionally complex. Our primary question was whether there was evidence for decompositional processing of the derived stems in Polish words – reflected in activation of the LH system that supports combinatorial grammatical processing – or for analysis of these stems as whole forms, reflected in cohort-like competition in the bilateral system.

Overall, we saw speech-driven lexical processing in bilateral temporal areas and in left inferior frontal regions (BA 47). The selective LIFG activation across the three conditions is consistent with earlier results for inflected forms in Polish (Szlachta et al., 2012), and arguably reflects the inflectional complexity of the case-marked nouns used here. A direct comparison between the three conditions revealed significant differences in the bilateral middle temporal gyri, driven by higher activation levels for opaque words compared to transparent and non-derived words. There were no differences between the three conditions in the LIFG. These results mirror the previously obtained English results in three critical respects.

First, there is the increase in activation for semantically opaque words in bilateral temporal areas, where this activity is linked to perceptual competition between the derived stem and its embedded base stem. This closely matches the results in English from Bozic et al. (2013), where semantically opaque forms (*archer*, *breadth*) led to increased activation in the bilateral processing system, compared to compositional (*bravely*) or simple forms (*giraffe*). The localisation of this activation was comparable across the two studies, peaking in the posterior middle and superior temporal gyri. These locations correspond to areas associated with lexical access (Hickok & Poeppel, 2007; Tyler & Marslen-Wilson, 2008). In Szlachta et al. (2012), a similar pattern of bilateral temporal activation was linked to perceptual competition between a word and an onset-embedded competitor (e.g., *kwitnie/kwit*). Similar factors dominate in accessing the lexical representations of Polish derived forms, reflecting cohort-based competition between whole-form representations of opaque derived stems and their onset-embedded stems and pseudostems.

The second major similarity is the absence of selective activation in the LH fronto-temporal system as a function of derivational complexity, even for semantically transparent complex words with regular suffixes like *pudełko*. This contrasts with the selective activation of the LH system for regular inflected English verbs (Bozic et al., 2010) and for Polish inflected verbs and nouns (Szlachta et al., 2012), all of which are argued to be assembled and disassembled online rather than stored as whole forms. Polish derived stems do not seem to activate the combinatorial LH processing system over and above simple stems, implying that they are not processed decompositionally in the same way as inflectional forms.

Thirdly, just as in the English study, the processing of transparent derived stems like *pudełko* does not generate increased activation relative to the simple stems. Competition effects in bilateral MTG are driven by the semantically opaque words, suggesting that the onset-embedded stems of transparently derived words do not function as cohort competitors to their full forms. This does not mean, however, that the transparent forms are decompositionally represented, since there is no evidence for the increased activation of LIFG associated with language-specific decompositional processes. Clahsen et al. (2003), based on behavioural data, suggest that semantically transparent derivations with productive affixes (as in the *pudełko* and *bravely* sets) are represented as whole forms, but with their morphological structure preserved. This provides a structured overlap between the lexical representations of full forms and of their embedded stems, reducing lexical competition for transparent derivations relative to opaque ones. This is reflected in similar activation levels for transparent and simple forms.

The close similarities, in conclusion, between the neural signatures of derivationally complex forms in English and Polish are consistent with broader inferences about how inflectional and derivational morphological processes interact with the proposed dual system approach. Derivational processes generate new stems in the language, which function neurocognitively as new lexical entities (or lexemes) that are primarily processed by the bilateral temporal systems that handle lexical access for morphologically simple forms. These derived stems enter into combination with inflectional morphemes to form surface phonological words, and it is the presence of these inflectional morphemes (optionally in languages like English, obligatorily in languages like Polish) that seems necessary to trigger selective involvement of the LH perisylvian system.

This analysis is broadly compatible with many aspects of existing linguistic, psycholinguistic and neuropsychological data (e.g. Hamilton & Coslett, 2008; Miceli & Caramazza, 1988). It is also consistent with recent neuroimaging results for the typologically different Finnish language (Leminen et al., 2011, 2013), which suggests wider cross-linguistic parallels. At the same time, the research raises as many questions as it answers. In particular, we need to understand how transparent derived stems combine the properties of whole-form representations with sufficient marking of internal decompositional structure to distinguish them neurocognitively from opaque forms.

## 4. Methods

### 4.1. Participants

Twenty-two right-handed native speakers of Polish were recruited for the study (eleven male, age range 18–36, mean age 27.3 years). They all had normal or corrected-to-normal vision, and no hearing, language or neurological impairments. Their length of stay in the UK did not exceed 6 years (having spent on average 30.8 months in the UK at the time of the study) and they all had at least 12 years of formal education in Poland. Two participants were later removed from the fMRI analyses due to movement artefacts.

### 4.2. Stimuli

There were three conditions of 60 items each (see Table 1) designed to vary derivational complexity and semantic transparency<sup>4</sup>: (1) Semantically transparent derived nouns; e.g., *pudełko*, 'little box'; (2) Semantically opaque derived nouns; e.g., *kanapka*, 'sandwich'; and (3) Simple nouns; e.g., *kapusta*, 'cabbage'. Transparent and opaque words were constructed using the same four affixes: diminutive *-k-*, *-anie/-enie* (equivalent of English gerund), *-ość* (equivalent of English *-ness*), and *-arz* (equivalent of English agentive *-er*). These four affixes were distributed equally (15 items each) across the two conditions. Simple words did not contain embedded stems or derivational suffixes.<sup>5</sup> All forms were inflectionally marked for nominative case, either by an overt affix or by a null morpheme (see Szlachta et al., 2012).

Semantic relatedness ratings between the stem and the derived word were obtained from a group of native Polish speakers using online questionnaires. The participants were asked to rate how similar pairs of words are in meaning on a scale 1–9, where 1 is 'not similar' and 9 is 'very similar'. The questionnaires also con-

<sup>4</sup> The experiment contained a further three inflected verb conditions, addressing a different set of questions. They are treated as fillers here, and the relevant comparisons will be reported elsewhere.

<sup>5</sup> The three sets of stimuli are listed in an Appendix in the Supplementary materials.

tained synonyms (e.g., *portfel*, ‘wallet’ – *portmonetka*, ‘purse’) and phonologically related words (e.g., *skrzydło*, ‘wing’ – *skrzynia*, ‘crate’) as fillers. Each word pair was rated by 15–19 participants. The average value was 7.23 for transparent words and 3.80 for opaque ( $t(118) = -13.56, p < .001$ ).

The three conditions were matched on length (number of phonemes and length of audio file) and form and lemma frequency using Match software (Van Casteren & Davis, 2007).

To match the total word length between conditions and control for the overall amount of processing they trigger, words with longer stems were selected for the simple condition. Average word length (with stem length in brackets) for the transparent, opaque and simple words was 737 (382), 740 (386) and 721 (721) ms respectively; with affix lengths of 355, 354 and 0 ms (data shown in Table S1 in the Supplementary materials). All words were recorded in a sound-proof room by a female native speaker of Polish. They were digitised at a sampling rate of 22 kHz with 16 bit conversion, and processed into separate files using Adobe Audition.

To isolate lexical processing from lower-level auditory processing we used an acoustic baseline called musical rain (MuR). MuR is closely matched to speech in terms of its temporal envelope and amount of energy, but does not produce the perception of speech (for details see Bozic et al., 2010; Uppenkamp, Johnsrude, Norris, Marslen-Wilson, & Patterson, 2006).

#### 4.3. Procedure

Participants listened attentively to the stimulus words, and performed an occasional one-back memory task (same/different), intended to maintain attention. Words were played every 3.4 s, and for 5% of the items a question appeared on the screen asking whether the sound they were currently hearing was the same as the previous one. They responded with a button-press (same = YES, different = NO). There were four blocks of 228 items each (with 12 questions per block), and 5 dummy items at the beginning of each block. There were short breaks between the blocks. The stimuli were delivered using in-house software and NNL Electrostatic headphones. A practice session was run outside the scanner prior to the experimental section.

#### 4.4. Data acquisition and analysis

Scanning was performed on a 3T Trio Siemens scanner at the MRC Cognition and Brain Sciences Unit, Cambridge. A fast-sparse gradient-echo EPI sequence was used to avoid scanner noise during stimulus presentation (repetition time [TR] = 3.4 s, acquisition time [TA] = 2 s, echo delay time [TE] = 30 ms, flip angle 78, matrix size  $64 \times 64$ , field of view [FOV] =  $192 \times 192$  mm, 32 oblique slices 3 mm thick, 0.75 mm gap). T1-weighted structural scans were obtained for anatomical localisation (3D MPRAGE sequence; TR = 2250 ms, TE = 2.99 ms, flip angle 9, FOV =  $256 \times 240 \times 192$  mm, matrix size =  $256 \times 240 \times 192$  mm, spatial resolution 1 mm isotropic).

Data analyses were performed in SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>). Preprocessing was done using Automated Analysis Recipes version 4 (<https://github.com/rhodricusack/automatic-analysis>) and included image realignment to correct for movement, segmentation, spatial normalisation to the MNI reference brain, and smoothing using 10 mm isotropic Gaussian kernel. Data were analysed using the general linear model with four blocks and 9 event types (3 noun conditions, 3 verb conditions, MRs, fillers, null). The neural response was modelled with the canonical haemodynamic response function. Motion regressors were included

to code for the movement effects. A high-pass filter with a 128 s cutoff was used to remove low-frequency noise. For group random effects analysis the contrast images were combined, and the results were thresholded at voxel level  $p < .001$  (uncorrected) and cluster level  $p < .05$  (corrected for multiple comparisons).

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.bandl.2013.09.001>.

## References

- Bertram, R., Schreuder, R., & Baayen, R. H. (2000). The balance of storage and computation in morphological processing: The role of word formation type, affixal homonymy, and productivity. *Journal of Experimental Psychology: Memory, Learning, and Cognition*, 26, 419–511.
- Binder, J. R., Frost, J. A., Hammeke, T. A., Cox, R. W., Rao, S. M., & Prieto, T. (1997). Human brain language areas identified by functional magnetic resonance imaging. *The Journal of Neuroscience*, 17, 353–362.
- Bozic, M., & Marslen-Wilson, W. D. (2010). Neurocognitive contexts for morphological complexity: Dissociating inflection and derivation. *Language and Linguistic Theory*, 11, 1063–1073.
- 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 USA*, 107, 17439–17444.
- Bozic, M., Tyler, L. K., Su, L., Wingfield, C., & Marslen-Wilson, W. D. (2013). Neurobiological systems for lexical representation and analysis in English. *Journal of Cognitive Neuroscience*, 25, 1678–1691.
- Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 257–294). San Diego, CA: Academic Press.
- Clahsen, H., Sonnenstuhl, I., & Blevins, J. P. (2003). Derivational morphology in the German mental lexicon: A dual mechanism account. In R. H. Baayen & R. Schreuder (Eds.), *Morphological structure in language processing* (pp. 125–155). Berlin: Mouton de Gruyter.
- Hamilton, A. C., & Coslett, H. B. (2008). Role of inflectional regularity and semantic transparency in reading morphologically complex words. *Neurocase*, 14, 347–368.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, 8, 393–402.
- Leminen, A., Lehtonen, M., Leminen, M., Nevalainen, P., Mäkelä, J., & Kujala, T. (2013). The role of attention in processing morphologically complex spoken words: An EEG/MEG study. *Frontiers in Human Neuroscience*. <http://dx.doi.org/10.3389/fnhum.2012.00353>.
- Leminen, A., Leminen, M., Lehtonen, M., Nevalainen, P., Ylinen, S., Kimppa, L., et al. (2011). Spatiotemporal dynamics of the processing of spoken inflected and derived words: A combined EEG and MEG study. *Frontiers in Human Neuroscience*, 5, 1–14.
- Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, 25, 71–102.
- Marslen-Wilson, W. D. (2007). Morphological processes in language comprehension. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics*. Oxford: OUP.
- Marslen-Wilson, W. D., & Tyler, L. K. (2007). Morphology, language and the brain: The decompositional substrate for language comprehension. *Philosophical Transactions of the Royal Society: Biological Sciences*, 362, 823–836.
- Marslen-Wilson, W. D., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review*, 101, 3–33.
- Miceli, G., & Caramazza, A. (1988). Dissociation of inflectional and derivational morphology. *Brain and Language*, 35, 24–65.
- Szlachta, Z., Bozic, M., Jelowicka, A., & Marslen-Wilson, W. D. (2012). Neurocognitive dimensions of lexical complexity in Polish. *Brain and Language*, 121, 219–225.
- Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology*, 57A, 745–765.
- Tyler, L. K., & Marslen-Wilson, W. D. (2008). Fronto-temporal brain systems supporting spoken language comprehension. *Philosophical Transactions of the Royal Society: Biological Sciences*, 363, 1037–1054.
- Uppenkamp, S., Johnsrude, I. S., Norris, D., Marslen-Wilson, W. D., & Patterson, R. D. (2006). Locating the initial stages of speech-sound processing in human temporal cortex. *NeuroImage*, 31, 1284–1296.
- Van Casteren, M., & Davis, M. H. (2007). Match: A program to assist in matching the conditions of factorial experiments. *Behavioural Research Methods*, 39, 973–978.
- Zhuang, J., Tyler, L. K., Randall, B., Stamatakis, E. A., & Marslen-Wilson, W. D. (2012). Optimally efficient neural systems for processing spoken language. *Cerebral Cortex*. <http://dx.doi.org/10.1093/cercor/bhs366>.