

Processing of zero-derived words in English: An fMRI investigation<sup>☆</sup>Christos Pliatsikas<sup>a,\*</sup>, Linda Wheeldon<sup>a</sup>, Aditi Lahiri<sup>b</sup>, Peter C. Hansen<sup>a</sup><sup>a</sup> School of Psychology, University of Birmingham, Birmingham B15 2TT, UK<sup>b</sup> Faculty of Linguistics, Philology and Phonetics, Centre for Linguistics and Philology, Walton Street, University of Oxford, Oxford OX1 2HG, UK

## ARTICLE INFO

## Article history:

Received 1 August 2013

Received in revised form

14 October 2013

Accepted 7 November 2013

Available online 15 November 2013

## Keywords:

Morphology

Functional MRI

Zero derivation

LIFG

## ABSTRACT

Derivational morphological processes allow us to create new words (e.g. *punish* (V) to noun (N) *punishment*) from base forms. The number of steps from the basic units to derived words often varies (e.g., *nationality* < *national* < *nation*: two-steps) and there is evidence that complex derivations cause more brain activity than simple ones (Meinzer, Lahiri, Flaisch, Hannemann, & Eulitz, 2009). However, all studies to date have investigated derivational processes in which morphological complexity is related to a change in surface form. It is therefore unclear whether the effects reported are attributable to underlying morphological complexity or to the processing of multiple surface morphemes. Here we report the first study to investigate morphological processing where derivational steps are not overtly marked (e.g., *bridge-N* > *bridge-V*) i.e., zero-derivation (Aronoff, 1980). We compared the processing of one-step (*soaking* < *soak-V*) and two-step (*bridging* < *bridge-V* < *bridge-N*) derivations together with monomorphemic control words (*grumble*) in an fMRI experiment. Participants were presented with derived forms of words (*soaking*, *bridging*) in a lexical decision task. Although the surface derived *-ing* forms can be contextually participles, gerunds, or even nouns, they are all derived from verbs since the suffix *-ing* can only be attached to verb roots. Crucially, the verb root is the basic form for the one-step words, whereas for the two-step words the verb root is zero derived from a basic noun. Significantly increased brain activity was observed for complex (one-step and two-step) versus simple (zero-step) forms in regions involved in morphological processing, such as the left inferior frontal gyrus (LIFG). Critically, activation was also more pronounced for two-step compared to one-step forms. Since both types of derived words have the same surface structure, our findings suggest that morphological processing is based on underlying morphological complexity, independent of overt affixation. This study is the first to provide evidence for the processing of zero derivation, and demonstrates that morphological processing cannot be reduced to surface form-based segmentation.

© 2013 The Authors. Published by Elsevier Ltd. All rights reserved.

## 1. Introduction

Derivation is the morphological process that leads to the creation of new words in a language. For example, the adjective *payable* is derived from the verb *pay* with the addition of the suffix *-able* that denotes possibility or necessity; similarly, the verb *reproduce* is derived from the verb *-produce* with the addition of the prefix *re-* which often denotes repetition or return to original state (Spencer, 1991). Such overt derivational processes (stem+affix) occur frequently in many languages. Another overt process involves not only addition, but also

deletion e.g., *enormity* is derived from *enormous* after truncating {-ous}, while in *pomposity*, the full form *pompous* remains. However, this is not the only derivational pattern that exists: in some cases, new words can be created without the addition of an affix, but simply with a change of syntactic class which does not affect the visual or phonological form of the word. For example, the verb *to bridge* is derived from the noun *the bridge* without the addition of any affix. This covert process of derivation is called **zero derivation** or conversion and is completely semantically compositional and very productive in English (Aronoff, 1980; Clark & Clark, 1979; Plank, 1981). Although processing of overt derivation in English has been well-studied and documented (Marslen-Wilson, Tyler, Waksler, & Older 1994; Rastle, Davis, & New 2004; Taft & Forster, 1975), to date there are no studies examining the processing of zero derivation, either at the behavioural or the neuronal level. This is the aim of the current study and is a critical issue since current theories of morphological processing reduce derivation to form-based processing of recognisable morphemes (Rastle et al., 2004; Taft & Forster, 1975) or to processing based on semantic regularities (Marslen-Wilson et al., 1994). Any experimental

<sup>☆</sup>This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

\* Corresponding author. Present address: School of Psychology, University of Kent, Canterbury, Kent CT2 7NZ, UK. Tel.: +44 1227 827411.

E-mail addresses: [c.pliatsikas@kent.ac.uk](mailto:c.pliatsikas@kent.ac.uk) (C. Pliatsikas), [l.wheeldon@bham.ac.uk](mailto:l.wheeldon@bham.ac.uk) (L. Wheeldon), [aditi.lahiri@ling-phil.ox.ac.uk](mailto:aditi.lahiri@ling-phil.ox.ac.uk) (A. Lahiri), [p.c.hansen@bham.ac.uk](mailto:p.c.hansen@bham.ac.uk) (P.C. Hansen).

evidence of the processing of zero derivation would suggest that morphological principles mediate the processing of complex words, and that processing is not reducible to form or semantics.

There is abundant evidence that morphologically complex derivations are automatically decomposed online (Marslen-Wilson et al., 1994; Rastle et al., 2004; Taft & Forster, 1975). Marslen-Wilson et al. (1994) used a cross-modal priming task to investigate the effect of morphologically related auditory primes on visual lexical decision. They showed that only semantically transparent morphological relationships resulted in priming; e.g., *punishment* and *punish* prime, but *casualty* and *casual* do not. Consequently, they claimed that morphological processing involves decomposition to a semantically determined base form. The masked priming technique has also shown that the recognition of a word is significantly facilitated if the word is preceded by a morphologically related prime (Frost, Deutsch, & Forster 2000; Rastle & Davis, 2008; Rastle et al., 2004). Rastle et al. (2004) used a masked priming lexical decision task to compare the processing of real morphological pairs (*cleaner-clean*) to pseudo-morphological pairs which contained morpheme-like chunks (*proper-prop*) and non-morphological pairs (*planet-plan*). The first word of each pair was presented very briefly preceded by a set of symbols and succeeded by the second word to which participants made a word/nonword lexical decision. They showed that both morphological as well as pseudo-morphological pairs primed, while non-morphological pairs did not. They proposed a model of morpho-orthographic processing in which morpheme-like constituents are automatically decomposed irrespective of real morphological relatedness (Rastle & Davis, 2008).

Evidence for the obligatory decomposition of morphologically complex words has also been provided by neuroimaging studies, which shows that decomposition processes manifest as increased brain activity (Devlin, Jamison, Matthews, and Gonnerman 2004; Gold & Rastle, 2007; Marslen-Wilson & Tyler, 2007). However, what remains to be determined is whether this decomposition is driven purely by overt morphemes or is governed by underlying morphological complexity. An early study of the neural bases of morphological decomposition used masked priming in an fMRI experiment (Devlin et al., 2004). Word pairs were presented which had overlapping orthography (*passive-pass*), semantics (*sofa-couch*), or morphology (*hunter-hunt*) and compared to unrelated control pairs (*award-munch*). Devlin et al. found significant reduction of brain activity in bilateral posterior angular gyrus for morphological compared to unrelated pairs, suggesting that some form of morphological processing occurred in this area. However, it was also shown that this reduction in activation completely overlapped with that observed for the orthographic and semantic conditions. They concluded that morphological processes are not independent but emerge as a combination of the processing of form and meaning. In contrast, Gold and Rastle (2007) demonstrated that this effect may only reflect the early and rapid decomposition of words with morpheme-like constituents (that has already been shown by behavioural experiments, such as Rastle et al. (2004)), rather than the processing of innate morphological relationships. In a masked-priming fMRI study, they identified a region in the anterior middle occipital gyrus (aMOG) with reduced brain activity for the processing of pseudo-morphological pairs (e.g. *archer-arch*). Critically, therefore, these studies revealed effects of morphological processing in visual areas, but showed no effects in brain areas that are traditionally linked to morphological processing. For instance, studies investigating the processing of inflections localise morphological processing in the LIFG (Marslen-Wilson & Tyler, 2007). Indeed, Dehaene et al. (2001, 2004) have suggested that effects from masked words may be limited to more posterior parts of the brain which underlie visual processing, and this task might therefore overlook morphological effects that take place in more anterior regions, including the LIFG.

More recently, the delayed repetition priming task has been used to isolate effects of morphological priming. In this task, the primes and targets are separated by a number of unrelated intervening stimuli and it has been shown to sustain morphological priming effects, while semantic and form effects deteriorate at longer lags (Bentin & Feldman, 1990; Napps & Fowler, 1987). Using this task, Bozic, Marslen-Wilson, Stamatakis, Davis, & Tyler (2007) tested prime-target morphological pairs which were either semantically transparent (*brave-bravely*) or opaque (*arch-archer*), and compared them to pairs that shared either form (*scan-scandal*) or meaning (*accuse-blame*). They found reduced activation of the LIFG for the presentation of the second member of morphologically related pairs, but not for pairs related in form or meaning. Moreover, there were no differences in the brain activity incurred by the processing of transparent versus opaque forms. Since the activated area did not overlap with areas that were shown to underlie processing of form or meaning related pairs, Bozic and colleagues concluded that the LIFG supports the processing of morphological structure which is not reducible merely to the processing of form and meaning.

Further evidence for the importance of the LIFG for morphological processes comes from Meinzer et al. (2009) who investigated the processing of multiply affixed forms. They predicted a relationship between derivational complexity and the degree of activation in the LIFG. They tested lexical decisions to morphologically complex German derived nouns in two conditions. In the one-step condition, nouns were derived directly from an adjective (e.g. *emsig* 'busy' > *Emsigkeit* 'being busy') or from a verb (e.g. *teilen* 'to separate' > *Teilung* 'separation'). In the two-step condition, nouns were either derived from a verb via an adjective (e.g. *trauern* 'to mourn' > *traurig* 'sad' > *Traurigkeit* 'sadness') or from an adjective via a verb (e.g., *eben* 'level' > *ebnen* 'to level' > *Ebnung* 'a level'). Since the same suffixes were used across the two conditions, any difference in brain activity could be attributed to the difference in the number of derivational steps that are necessary to derive back to the base of each of these forms, i.e., the depth of the derivation. Meinzer et al. reported increased activity in the LIFG for two-step versus one-step nouns, which was accompanied by increased activity in bilateral superior temporal areas and occipital areas. The authors interpreted their findings as indicative of the obligatory decomposition of all constituents of multi-affixed forms.

### 1.1. The present study

The emerging picture of the visual processing of complex words is one in which there is an early stripping of affix-like constituents localised in posterior brain areas followed by the processing of morphological structure localised in the LIFG. However, it remains unclear whether this morphological processing is driven by the processing of overt affixes or by underlying morphological composition. The current study aims to address this issue. The results of Meinzer et al. could indeed be attributed to the obligatory morphological derivation of a complex form back to its base. Conversely, these findings may simply reflect the processing of the extra affix in two-step compared to one-step forms. In this sense, it is hard to distinguish between surface decomposition of stems and affixes and underlying derivational complexity. We therefore tested the processing of zero derivation forms in English within the context of a lexical decision fMRI experiment. Similar to Meinzer et al., we compared two-step to one-step derived forms. Critically, in our experiment, the two-step forms are derived from their root via intermediate zero-derived forms (e.g. *bridge-N* > *bridge-V* > *bridging*), whereas the one-step forms are directly derived from their root (*write-V* > *writing*). The intermediate step in the two-step derivation is crucial because, although the base form of *bridge* is a noun, the suffix *-ing* can only attach to verbs.

So *-ing* can be added to the verb form *bridge* to form *bridging* but cannot, for example, add to *desk*, which does not (at the moment at least) have a transparent verb form. Critically however, although the suffix {-ing} has more than one function, it is productively used only with verbs; for example, (i) gerunds (*Killing tigers is prohibited*), (ii) active participle (*He is killing a tiger*), (iii) action nominalisations (*The killing of a tiger*), (iv) adjectives (*man-eating tigers*). One could argue that instances of noun-deriving {-ing}s also occur, such as *bedding* ‘articles which are used on the bed’ or *legging* ‘covering for the leg’, but these are extremely marginal and such forms were not used in the experiment.

The suffixed forms we tested therefore have the same surface structure (stem + *-ing*) but differ in the complexity of their underlying derivations. This contrast allows us to investigate the effect of derivational depth independently from surface complexity and to test which brain areas carry out such covert morphological operations. Additionally, we aimed to ensure that any effects are due to the morphological complexity of our derivations, and not to factors which are known to affect word recognition, such as concreteness, frequency, orthographic and phonological properties. For this reason, we collected norming data on a number of critical measures and matched our materials as closely as possible on a wide range of lexical properties. We also conducted separate analyses, including these factors as covariates, in order to tease apart any effects not related to morphology.

On the basis of previous findings (Devlin et al., 2004; Gold & Rastle, 2007), we predicted that the morphologically complex forms (in comparison to morphologically simple forms), would engage the bilateral occipital regions that have been shown to subserve automatic decomposition of complex forms. Moreover, if the Meinzer et al. results are due to derivational complexity, we predicted that our two-step forms, although not overtly affixed, should nevertheless elicit more brain activity in the LIFG than our one-step forms. If, on the other hand, morphological processing is only reducible to surface complexity, then increased LIFG activation for our zero-derived two-step forms should not be observed.

## 2. Methods

### 2.1. Participants

Twenty-two undergraduate or postgraduate students participated in the experiment (male/female, mean age: 20.4 y, SD: 2.96). They were all native speakers of British English, had normal or corrected-to-normal vision, and were strongly right-handed, as assessed by a handedness inventory (Annett, 1972). All participants also scored highly on three reading tasks: the TOWRE (Torgesen, 1999) which tested their ability to read real printed words (mean score: 85%, SD: 0.11%) and pronounceable printed nonwords (mean score: 92%, SD: 0.06%), and the TIWRE (Reynolds & Kamphaus, 2007) which tested their reading abilities for real irregular printed words (mean score: 97%, SD: 0.02%). Participants were recruited from within the University of Birmingham, and were awarded with course credit. Based on the fMRI data preprocessing, one participant's fMRI data set, and two of the six fMRI blocks from another participant, were excluded from further analysis due to excessive head movement, defined as frequent displacement over 3 mm from the reference scan image. Ethical consent for this study was approved by the University of Birmingham Central Ethics Committee (Ethics code ERN\_11-0429AP8).

### 2.2. Materials

The experimental conditions consisted of two sets of 30 disyllabic derived *-ing* verbs, with initial stress: (a) one-step verbs (e.g. *soaking*), (b) two-step verbs (e.g. *bridging*). We also constructed a set of 30 monomorphemic disyllabic control verbs with initial stress. It was important that the control verbs had a similar degree of orthographic and phonological overlap to the experimental items in their final segments. Very few phoneme sequences occur frequently as verb endings without involving morphological structure. Our control verbs therefore all shared an ending of *-le* (e.g. *grumble*), which does not function as a morpheme in English. The word lists appear in Table 1.

The derived verbs all had stems that were either basic nouns or basic verbs. Based on the linguistics literature, we hypothesised that all one-step verbs had

stem meanings that would be judged to be action based and therefore basic verbs (e.g., *soak*), whereas all two-step verbs had stems that would be judged to be object based and therefore basic nouns (e.g., *bridge*). These predictions were confirmed by a rating pre-test, which was administered to a different group of 20 students. It included all the stems from our experimental items, and the participants were required to indicate how much they thought each stem referred to an action, on a 1–9 scale (9 = only referring to an action). The Action ratings appear in Table 2, and revealed that one-step verb stems were judged to refer to an action significantly more than the two-step stems.

In addition the one- and two-step verb sets were matched for a number of factors that can affect lexical recognition (see Table 2). These included word-length (in terms of both number of letters and phonemes), bigram and trigram frequency, and orthographic, phonological and morphological neighbourhood size taken from the CELEX database (Baayen, 1995), using N-Watch software (Davis, 2005). The whole word forms were also matched on a number of frequency measures (CELEX written and spoken counts, and the Bank of English (Järvinen, 1994) spoken counts per million). Stem surface frequencies for the experimental words were also matched using the Bank of English database. Additionally, we extracted the frequencies of the verb stem forms (e.g. *run-V*) and the noun stem forms (e.g. *run-N*), which, unsurprisingly, differed across experimental word sets. Verb stem frequency was higher for the one-steps than the two steps and this difference approached significance. Noun stem frequency showed a significant effect in the opposite direction, with a higher frequency for the two-steps than the one-steps. Interestingly, for both one- and two step verb stems, the base noun frequencies were significantly higher than their base verb frequencies (One-steps: Verb stems 1.1, Noun stems 2.2:  $t(29) = 2.088$ ,  $p = 0.046$ ; Two-steps: Verb stems 0.5, Noun stems 6.1,  $t(29) = 4.154$ ,  $p < 0.001$ ). This is perhaps due to the frequent use of one-step stem forms in noun contexts e.g., “I’m going for a run”.

The word sets were also matched on concreteness and imageability ratings collected from 20 participants using rating questionnaires. The questionnaires consisted of the experimental items along with filler words, and were identical for both rating tasks, apart from the instructions. For concreteness ratings, participants were asked to rate each word according to the extent to which they referred to concrete objects on a scale from 1 to 9, when 9 represented “very concrete”. For imageability ratings, participants were asked to rate the same words according to how well they elicited some sensory experience (mental image, smell, etc.), where 9 represented “very imageable”. None of the participants who took part in the rating pre-tests completed either questionnaires or the Action pre-test, nor were they tested in the main experiment. Finally, we were concerned that our word sets should not differ in terms of general difficulty when they are included in a lexical decision task. Therefore a different group of 20 students filled out an acceptability

**Table 1**  
List of experimental stimuli.

One-step verbs	Two-step verbs	Control verbs
Biting	Bleeping	Baffle
Blending	Boating	Chortle
Blurring	Bolting	Crinkle
Blushing	Boozing	Dawdle
Bonding	Bossing	Dazzle
Booming	Bridging	Dribble
Breaching	Brushing	Dwindle
Brewing	Camping	Fumble
Browsing	Cloaking	Gamble
Bruising	Fibbing	Grumble
Bumping	Greasing	Gurgle
Buzzing	Herdling	Huddle
Churning	Plugging	Hustle
Conning	Potting	Jostle
Diving	Rafting	Jumble
Dreading	Roofing	Mingle
Flirting	Scouting	Ruffle
Glaring	Seeding	Rustle
Grabbing	Shading	Scramble
Hoaxing	Shelling	Scribble
Quizzing	Shipping	Shuffle
Scaring	Sleeting	Strangle
Scorning	Sleighting	Stumble
Slitting	Snacking	Swindle
Snoozing	Spicing	Tremble
Soaking	Tagging	Twinkle
Tanning	Thumbing	Wangle
Thirsting	Trimming	Wobble
Trekking	Waxing	Wrangle
Tricking	Wheeling	Wriggle



**Table 2**  
Experimental stimuli matching.

	One-step	Two-step	<i>p</i>	Control	<i>p</i>
<b>Semantic counts</b>					
Acceptability judgement (%)	95.3 (14)	89.2 (12)	0.19	91.5 (13)	0.91
Concreteness rating (1–9)	4.8 (1.1)	5.3 (0.9)	0.23	4.8 (1.0)	0.11
Imageability rating (1–9)	5.5 (1.6)	5.6 (1.1)	0.75	5.3 (1.5)	0.69
Stem action rating (1–9)	6.5 (1.6)	3.7 (1.9)	0.000	n/a	
<b>Frequency counts</b>					
BoE surface spoken	0.7 (0.6)	0.5 (0.6)	0.47	0.6 (0.7)	0.54
BoE stem	7.7 (8.0)	10.0 (10)	0.40	n/a	
BoE verb base	1.1 (1.7)	0.5 (0.8)	0.08	n/a	
BoE noun base	2.2 (2.5)	6.1 (7.5)	0.009	n/a	
CELEX written	2.0 (2.1)	1.6 (2.4)	0.76	1.9 (1.5)	0.72
CELEX spoken	0.6 (0.9)	0.4 (0.6)	0.50	0.8 (1.3)	0.24
CELEX bigram	2.7 (0.3)	2.8 (0.2)	0.10	2.5 (0.3)	0.000
CELEX trigram	1.9 (0.4)	2.0 (0.3)	0.24	1.5 (0.5)	0.000
<b>Neighbourhood density counts</b>					
Orthographic	2.1 (2.5)	2.8 (2.1)	0.33	1.8 (1.9)	0.19
Phonological	5.5 (4.3)	6.7 (4.9)	0.58	4.9 (2.3)	0.21
Morphological	7.3 (2.6)	8.4 (3.4)	0.40	7.2 (2.4)	0.21
<b>Length counts</b>					
Phonemes	5.8 (0.7)	5.5 (0.6)	0.17	6.0 (1.0)	0.051
Letters	7.6 (0.7)	7.5 (0.6)	0.64	6.6 (0.7)	0.000

Stem frequency is the summed frequencies of all words in which the stem occurs. The base noun and verb frequencies are the frequency of occurrence of the bare noun and verb stems. This is because lexical databases do not distinguish between noun and verb stems in complex forms e.g. “bounces” as the plural of the noun bounce or the third person of the verb to bounce. These numbers therefore provide only an indication of the frequency of occurrence of the noun and verb forms of the stems.

judgment pre-test, where they had to judge whether each of the experimental words was a real word in English or not. As can be seen in Table 2, the one- and two-step verb sets did not differ in acceptability.

The same measures were obtained for the control verbs, except for the stem frequency counts and the stem action ratings, which did not apply to these items. Analyses of variance were conducted to test whether these items differed from the experimental sets and the means and *p* values are also shown in Table 2. The control verbs differed from the experimental sets only in the length measures. The differences were extremely small but nevertheless significant. The control forms were significantly shorter by less than one letter and also significantly longer by less than one phoneme, compared to the experimental lists. Additionally, they had slightly but significantly smaller bigram and trigram frequencies. These differences were a consequence of the limited choice of non-morphological endings and their lack of productivity. To determine the contribution of these and the other factors on our pattern of results they were added as covariates in separate models reported in Section 3.

In order to obscure the purpose of the experiment we also included 24 filler words ending in *-ity*, 24 filler words ending in *-ness*, and a further 18 filler words ending in *-ing* or *-le*, which did not have the characteristics required for the matched experimental word sets. There were therefore 108 unique actual words presented. For the purposes of the lexical decision task we also included an equal number (108) of various types of grammatically invalid but pronounceable non-words: (a) 48 *-ing* forms, created by the affixation of adjectives with the suffix *-ing* (e.g. *\*fuling*), (b) 24 *-le* forms (e.g. *\*jeggles*), (c) 18 *-ness* forms, created by the affixation of verbs with the *-ness* suffix (e.g. *\*ignoreness*) and (d) 18 *-ity* forms, created by the affixation of nouns with the *-ity* suffix (e.g. *\*venomity*). The chosen form of these nonwords made them relatively more plausible and the lexical judgment therefore more demanding. All fillers and nonwords were also included in the acceptability judgment pre-test. All nonwords scored very low in the pre-test, in terms of the percentage of the participants that judged them to be real words. The filler words and the nonwords were similar in length to the experimental and control words.

Two experimental runs were created, each containing all the experimental words and half of the fillers and the nonwords, resulting in 210 items per run. Participants were asked to make a lexical decision in response to each trial presentation but to press a button in response only to nonwords. All participants undertook both experimental runs, resulting in the experimental items being presented twice, in order to maximise the statistical power of our experiment. The materials within each run were pseudorandomised. Each run was divided into three blocks with an equal number of trials, and the blocks were permuted in three different ways, creating three experimental versions of six blocks each, across which the participants were evenly distributed. Finally, a practice block of 36 items was created, including different word and nonword forms of all types, in a similar ratio to the experimental blocks.

### 2.3. Design

An rapid-presentation event-related design was selected for this study, with variable inter-trial intervals (ITIs) (Dale, 1999; Dale, Greve, & Burock, 1999; Josephs & Henson, 1999). Each of the six experimental blocks had a unique event-related design in terms of the ITIs and the stimuli permutations. The designs were made through the optseq2 software (<http://www.freesurfer.net/optseq/>), calculated on the basis of Repetition Time (TR)=2.5 s, 210 brain volumes and 5 experimental conditions. Each trial was shown for 500 ms, followed by the variable ITI.

### 2.4. Procedure

Participants were each scheduled for two visits. During Visit 1 they completed the handedness questionnaire and the MRI screening form, and were pre-tested with the reading tasks. Upon successful completion of these tasks, they were invited to participate in the fMRI experiment. During Visit 2 participants were scanned during the experimental session, following a practice session which took place outside of the scanner. The experiment was presented using Presentation<sup>®</sup> software ([www.neurobs.com](http://www.neurobs.com)) via back projection. Stimuli were presented in black letters over a white background between two vertical red lines that stayed on the screen at all times and acted as a fixation point. Subjects were given a button response box to their left hand. They were instructed to press a button with their index finger as fast as they could whenever they saw a nonword, and not to respond when they saw a real word. This was because we were not concerned with their response times to real words, and we wanted to avoid any contamination to our brain activity data for the trial of interest, which could be caused by motor activation underlying the button press per se. The experimental session included six echo planar imaging (EPI) blood oxygen level-dependent (BOLD) scans, one for each of the experimental blocks, separated by short breaks. These were followed by acquisition of an anatomical T1-weighted image of the participant's head. The scanning session lasted for about 1 h 10 min.

### 2.5. fMRI data acquisition

Scanning was carried out on a 3T Phillips Achieva scanner situated at the Birmingham University Imaging Centre (BUIC). An 8-channel head coil was used. Functional images were acquired by using a T2\*-weighted gradient-echo EPI sequence with 42 axial slices of resolution  $3 \times 3 \times 3 \text{ mm}^3$ , acquired from bottom to top (TE: 35 ms, TR: 2.5 s, flip angle: 79°, FOV:  $240 \times 240 \text{ mm}$ , in-plane matrix resolution:  $80 \times 80$ , SENSE factor: 2). High resolution T1-weighted gradient echo anatomical images were collected with  $175 \times 1 \text{ mm}$  sagittal slices (TE: 3.7 ms, FOV:  $288 \times 232 \times 175 \text{ mm}$ ).

### 2.6. fMRI data analysis

Raw structural and functional data were converted from Phillips PAR/REC format into NIFTI format. All data processing was carried out using FEAT v5.98, part of FSL v4.1.8 (Smith et al., 2004; Woolrich et al., 2009).

Each of the six functional scans per participant were analysed separately. Non-brain tissue was removed from the high resolution anatomical images using BET (Smith, 2002). The functional data were motion-corrected using MCFLIRT (Jenkinson, Bannister, Brady, & Smith, 2002), which applied rigid body transformations to the acquired images, using the middle image of the time series as the template image, in order to optimise the alignment of the functional images. The data were also slice-time corrected using Fourier-space time-series phase-shifting in order to correct for temporal offsets in the acquisition of the images. Images were spatially smoothed using a Gaussian kernel (FWHM=5 mm) in order to reduce spatial noise from the data. Grand-mean intensity normalisation of the 4D datasets was applied, as well as high pass temporal filtering with Gaussian-weighted least-squares straight line fitting (sigma=45 s), in order to remove low-frequency artefacts.

Data from each experimental block were analysed using a general linear model, with the three experimental conditions (one-step, two-step, control), modelled as three separate events, and the fillers and the nonwords modelled as two additional events of no interest. A sixth event of no interest was added, modelling the button presses made to the nonwords by the participants, as recorded by Presentation. This event included the onset of the button press and a notional duration, defined as 100 ms, for all button presses. Stimuli timecourses that modelled the onset and duration of each condition event were convolved with a Double-Gamma HRF. Temporal filtering was applied to the model and temporal derivatives were added as separate regressors, in order for the model to better fit the time course of the actual data acquisition. Time-series statistical analysis was carried out using FILM with local autocorrelation correction (Woolrich, Ripley, Brady, & Smith 2001). Finally, the motion parameters generated by MCFLIRT were added to the model as separate regressors of no interest, in order to correct for any residual artefacts caused by motion (Johnstone et al., 2006).

In order to investigate how morphological processing is related to depth of derivation, the (two-step+one-step) > control and two-step > one-step contrasts

**Table 3**  
Differential activations for the contrasts of interest.

Contrast	Structures	Hemi	Voxel count	PeakZ	MNI coordinates		
					x	y	z
(One-step+two-step) > control	Occipital fusiform gyrus	L	2001	4.03	−18.5	−83.4	7.28
	Inferior frontal gyrus	L	1645	4.03	−48.8	23.1	4.62
	Occipital fusiform gyrus	R	1201	3.78	14.9	−80.2	−3.02
	Inferior temporo-occipital gyrus	L	526	3.39	−58.5	−50.5	−12.9
Two-step > one-step	Inferior frontal gyrus	L	1143	4.01	−47.2	22.4	5.37

were calculated per voxel across the whole brain. The statistical images were registered to the high resolution image of the participant using a 12 DOF affine registration in FLIRT (Jenkinson et al., 2002; Jenkinson & Smith, 2001). The high resolution image was subsequently registered to the 152-brain T1-weighted MNI template with the use of FNIRT non-linear registration (warp resolution = 10 mm) (Andersson, Jenkinson, & Smith 2007a, 2007b).

At second level analysis, all experimental blocks for all subjects were added in a single model, where subjects were modelled as separate events. Each subject was modelled as a contrast, created by the combination of the experimental blocks for each participant using a fixed-effects model in FLAME (Beckmann, Jenkinson, & Smith, 2003; Woolrich, 2008; Woolrich, Behrens, Beckmann, Jenkinson, & Smith, 2004), by forcing the random effects variance per participant to zero. This modelling procedure was chosen because it accounts for the variability among sessions, but also among subjects too.

At third level analysis, each contrast of interest was analysed separately using a mixed-effects model in FLAME stage 1 and stage 2 (Beckmann et al., 2003; Woolrich et al., 2004; Woolrich, 2008), where the second level contrast images from each participant were input as a single event. The resulting statistic image was cluster thresholded in a two-step procedure determined by a voxel threshold of  $Z > 2.3$  and a corrected cluster significant threshold of  $p < 0.05$  (Worsley, 2001).

In order to ensure that other linguistic properties of the word lists had not affected the pattern of our MRI results, the third-level analysis was additionally run with the incorporation the linguistic features of our stimuli (as they appear in Table 2) as covariates. This included an additional analysis with the difference between the noun and the stem frequencies as a covariate.

### 3. Results

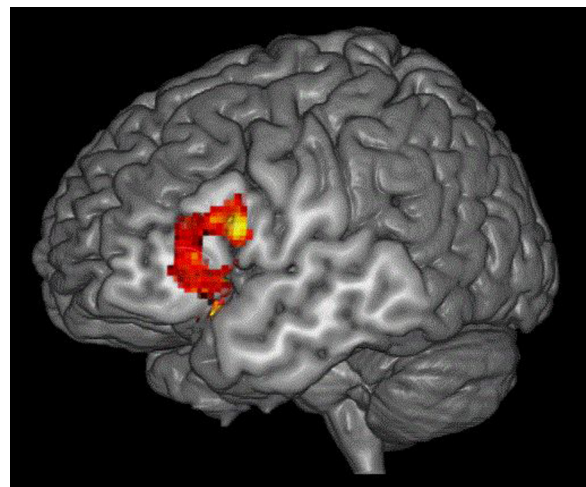
#### 3.1. Behavioural data

Reaction time data were only available for nonword trials as participants were instructed not to respond to the stimuli they judged to be real words. Participants were highly accurate in detecting the nonwords (mean accuracy: 91.9%, SD: 6.7%). Their average reaction time for the correct trials was 1.07 s. (SD: 0.11).

#### 3.2. fMRI data

We first examined increases or decreases in brain activity caused by the one-step and two-step experimental conditions collapsed together, compared to the control condition. This contrast illustrates brain activity that is specific to morphologically complex words, as described by previous research. Increased brain activity for complex words was found in several areas, most importantly in the left IFG, pars opercularis, which has already been shown to underlie processing of morphologically complex words (Meinzer et al., 2009). Additionally, increased activation was revealed in the left temporo-occipital gyrus, reflecting the rapid form-based decomposition of complex forms (Gold & Rastle, 2007) and the occipital fusiform gyrus, bilaterally.<sup>1</sup> Table 3 illustrates the significant areas of activation for this contrast.

<sup>1</sup> It is possible that this bilateral activation for complex vs. simple forms, not reported in any of the previous studies, is not due to any decomposition taking place, but due to the physical properties of the forms; indeed, our complex forms are significantly longer by one letter compared to the simple forms, as it appears in Table 2. However, this is a very small difference in length; therefore its effect on



**Fig. 1.** Differential activation for two-step versus one-step verbs in the LIFG.

More importantly for this investigation the two-step > one-step contrast was examined, in order to indicate whether areas of the brain specifically process implicit morphological structure. This contrast was masked inclusively with the results from the (one-step+two-step) > control contrast, as they appear in Table 3, in order to check for any differences between the two conditions within those areas that underlie processing of complex forms. The analysis revealed a large cluster of activation in the left IFG, pars opercularis and triangularis, extending to the left orbital frontal gyrus and the left insula. The differential activation for two-step versus one-step conditions is reported in Table 3 and the activation pattern of the two-step > one-step contrast in the LIFG is shown in Fig. 1, overlaid onto a standard template brain for illustrative purposes. Fig. 2 also shows the time course of the BOLD response in the activated area for the three main experimental conditions.

#### 3.3. Covariate analysis

Separate third-level analyses were re-run on the masked 2step > 1step contrast, where each of the parameters in Table 2 was added as a covariate, with an additional analysis having the noun vs. verb stem frequency difference as a covariate. This was done in order to ensure that the observed differences for this contrast were not due to lexical or physical properties of the two lists. None of these covariates produced significant activation. Importantly, this included features in which the experimental lists differed significantly, such as the action ratings. Based on that,

(footnote continued)

occipital activation cannot be readily determined, especially with our experimental materials.

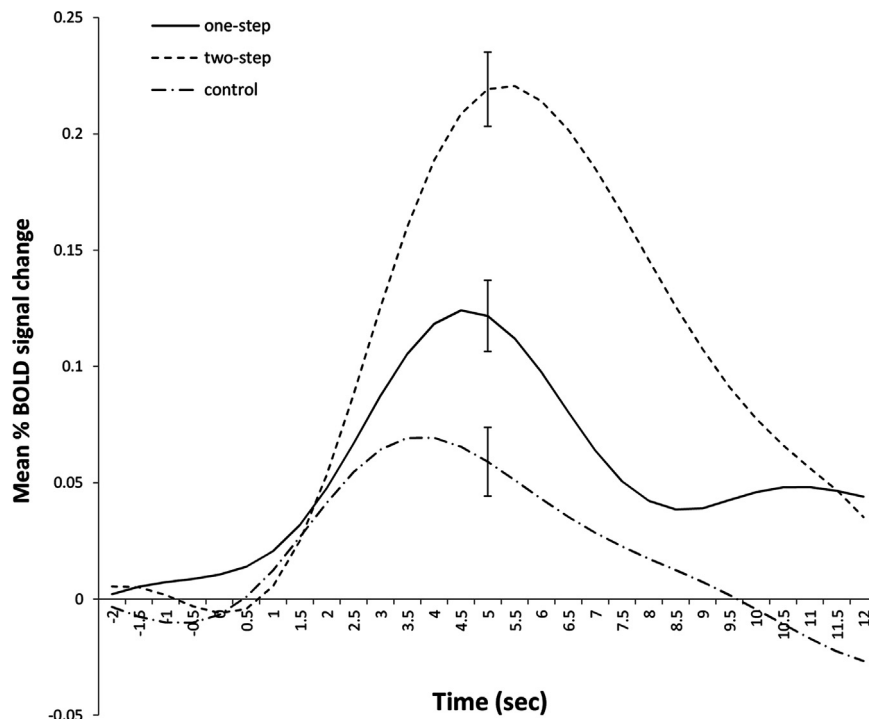


Fig. 2. Mean % BOLD signal change for two-step, one-step and control verbs in the LIFG ( $\pm$  SEM,  $N=21$ ).

we believe that none of the observed effects can be explained by the lexical or physical properties of the lists.

#### 4. Discussion

This study aimed to investigate the neural bases of derivational processing in native speakers of English. In particular, we wished to determine whether morphological processing is limited to the decomposition of overt affixes or also includes the processing of covert derivational complexity. The evidence available to date has shown that brain activity increases as a function of the number of overt morphemes (Meinzer et al., 2009). Our fMRI study tested zero-derived forms in which the derivation is not associated with any surface change (e.g., *bridge-N* > *bridge-V*). We report here two main findings; first, morphologically complex forms (e.g. *soaking* one-step, *bridging* two-step) engage a posterior-to-anterior brain network compared to monomorphemic words (e.g., *grumble*). Second, where the first step is zero-derivation, two-step forms (e.g., *bridging*) generate more activity in the LIFG than one-step forms (e.g. *soaking*).

Our results are, in a number ways, compatible with the available neuroimaging findings on derivational processing. For example, the observed activation of the occipital lobe for complex versus simple forms has been reported previously (Devlin et al., 2004; Gold & Rastle, 2007). Our data is consistent with the claim that this particular region underlies the early recognition of surface morpheme-like units in the visual input. Critically, we observed no differential activation in this area for our two-step versus one-step words, which do not differ in their surface form properties (i.e., stem+*-ing*). Similarly, the activation of LIFG for the processing of derived compared to simple words has already been reported in a delayed priming task, which showed no comparable effects of form or meaning (Bozic et al., 2007). Therefore, our data also confirm the importance of this area for the processing of morphological structure.

Our results also revealed a similar pattern of activation to that reported by Meinzer et al. (2009), who investigated the effects of derivational depth in the LIFG. They found that two-step derivations caused increased activity in the LIFG, compared to one-step

derivations. Based on this finding, Meinzer and colleagues argued that the processing of complex words involves an obligatory derivation back to the base form and that the processing in the LIFG is primarily related to derivational depth. Our study also tested the processing of one-step and two-step derivations and yielded significant effects of derivational depth in the LIFG. Our data therefore suggest a similar explanation. We propose that speakers know the rules of combination for the morphemes of their language e.g., one can add the suffix *-ing* only to verbs. The derivational process is entirely compositional and predictable. Even when the surface derived form like *watering* can be contextually employed as an active nominal (*Her watering the plants was successful*), the compositional process is *water-Noun* > *water-Verb*, *water-ing - participle* > *watering Noun*. It follows therefore that stripping an *-ing* suffix must necessarily activate a verb stem. For one-step verbs, the verb stem is the base form and no further stem access is required. For two-step forms the noun stem is the base form and a successful word recognition process would also involve the access of this form.

However, our findings do not, of course, provide direct information concerning the nature of the processes underlying the differences in brain activity we have observed in the LIFG. An alternative but related possibility is that the high activation in the LIFG for two-step verbs is not due to the additional derivational step we postulate but is attributable to a process of disambiguation or competition between the noun and verb stems. Following the stripping of the *-ing* suffix, the remaining stem might activate both its noun and verb forms and the degree of activation could be a function of their frequency of use. In other words, stem frequency could determine which form is accessed first. Average frequencies for noun stems were higher than for verb stems for both our one-step and two-step items. However the difference in stem frequencies was much larger for two-steps than for the one-steps. If a correct lexical decision requires the association of the suffix with a verb form, then a higher frequency noun stem might slow the retrieval of the required verb stem. Based on stem frequencies this competition would be greater for the two-steps than the one-steps, leading to increased neural activity.

This explanation is similar to our derivational explanation in that it requires that noun and verb stems are represented in the



mental lexicon, as is knowledge of morphological processes i.e., -ing is attached only to verbs. It also assumes that resolution of the conflict is based on morphological knowledge i.e., *bridging* is a word because the verb *bridge* exists, whereas *dooring* is not a word because the verb *door* does not exist. However, the critical difference lies with the order of stem activation, i.e. whether the order is determined by morphological rules or by the frequency-weighted activation of both stems. Although our data do not allow us to decide conclusively between these alternative explanations, the frequency based competition explanation is unsupported by the results of the analyses of covariance, which show that neither noun stem frequency nor the difference between noun and verb stem frequencies significantly contribute to the critical difference in the LIFG. Further experimentation is nevertheless required and behavioural experiments testing for competition effects are underway.

Recall that the central aim of our experiment was to test for morphological processes that could not be reduced to surface decomposition. The critical difference between our study and that of Meinzer et al. is in the relationship between derivational and surface complexity. Our two-step derived word forms were only covertly more complex than the one-step forms, as both forms had the same number of surface morphemes. Our findings therefore provide a strong challenge to the assumption that morphological processes are reducible to the decomposition of surface constituents. Instead, we argue that derivational complexity mediates online morphological processing even in the absence of any related change in surface form. This is an important finding, as it demonstrates, for the first time, the neurological reality of covert morphological complexity.

What is now apparent is that an accurate model of morphological processing must go beyond the mapping between surface morphemes and a semantically determined morphological lexicon. After all, morphological processes in many languages are associated with structural reductions or changes or, as we have shown, no surface changes at all. It is important that we investigate the underlying relationships between such disparate forms as they represent the speakers' understanding of the morphological structure of their language and must determine their recognition and generation of both known and novel words.

## Acknowledgements

The study was funded by the ESRC (Project code: RES-000-22-4222). We are grateful for the very useful comments from the reviewers. We would also like to thank Debra Malpass for her contribution to the design of the study and the creation of the stimulus lists.

## References

- Andersson, J. L. R., Jenkinson, M., & Smith, S. M. (2007a). Non-linear optimisation. *fMRIB Analysis Group Technical Reports: TR07JA02*. Available from [www.fmrib.ox.ac.uk/analysis/techrep](http://www.fmrib.ox.ac.uk/analysis/techrep) Accessed on 01.08.13.
- Andersson, J. L. R., Jenkinson, M., & Smith, S. M. (2007b). Non-linear registration, aka spatial normalization. *fMRIB Analysis Group Technical Reports: TR07JA2*. Available from [www.fmrib.ox.ac.uk/analysis/techrep](http://www.fmrib.ox.ac.uk/analysis/techrep) Accessed 01.08.13.
- Annett, M. (1972). The distribution of manual asymmetry. *British Journal of Psychology*, 63(3), 343–358.
- Aronoff, M. (1980). Contextuals. *Language*, 56(4), 744–758.
- Baayen, R. H. (1995). *The CELEX lexical database (release 2)*. Philadelphia, PA: Linguistic Data.
- Beckmann, C., Jenkinson, M., & Smith, S. M. (2003). General multi-level linear modelling for group analysis in FMRI. *NeuroImage*, 20, 1052–1063.
- Bentin, S., & Feldman, L. B. (1990). The contribution of morphological and semantic relatedness to repetition priming at short and long lags: Evidence from Hebrew. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 42(4), 693–711.
- Bozic, M., Marslen-Wilson, W. D., Stamatakis, E. A., Davis, M. H., & Tyler, L. K. (2007). Differentiating morphology, form, and meaning: Neural correlates of morphological complexity. *Journal of Cognitive Neuroscience*, 19(9), 1464–1475.
- Clark, E. V., & Clark, H. H. (1979). When nouns surface as verbs. *Language*, 55(4), 767–811.
- Dale, A. (1999). Optimal experimental design for event-related fMRI. *Human Brain Mapping*, 8(2–3), 109–114.
- Dale, A., Greve, D., & Burock, M. (1999). Optimal stimulus sequences for event-related fMRI. *NeuroImage*, 9, 533.
- Davis, C. J. (2005). N-Watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, 37(1), 65–70.
- Dehaene, S., Jobert, A., Naccache, L., Ciuciu, P., Poline, J. B., Le Bihan, D., & Cohen, L. (2004). Letter binding and invariant recognition of masked words. *Psychological Science*, 15(5), 307–313.
- Dehaene, S., Naccache, L., Cohen, L., Le Bihan, D., Mangin, J. F., Poline, J. B., & Rivière, D. (2001). Cerebral mechanisms of word masking and unconscious repetition priming. *Nature Neuroscience*, 4(7), 752–758.
- Devlin, J. T., Jamison, H. L., Matthews, P. M., & Gonnerman, L. M. (2004). Morphology and the internal structure of words. *Proceedings of the National Academy of Sciences of the United States of America*, 101(41), 14984–14988.
- Frost, R., Deutsch, A., & Forster, K. I. (2000). Decomposing morphologically complex words in a nonlinear morphology. *Journal of Experimental Psychology: Learning, memory and Cognition*, 26(3), 751–765.
- Gold, B. T., & Rastle, K. (2007). Neural correlates of morphological decomposition during visual word recognition. *Journal of Cognitive Neuroscience*, 19(12), 1983–1993.
- Järvinen, T. (1994). Annotating 200 million words: the Bank of English project. In *Proceedings of the 15th International Conference on Computational Linguistics COLING 94* (pp. 565–568). Kyoto: Association for Computational Linguistics.
- Jenkinson, M., Bannister, P. R., Brady, M., & Smith, S. M. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage*, 17(2), 825–841.
- Jenkinson, M., & Smith, S. M. (2001). A global optimisation method for robust affine registration of brain images. *Medical Image Analysis*, 5(2), 143–156.
- Johnstone, T., Ores Walsh, K. S., Greischar, L. L., Alexander, A. L., Fox, A. S., Davidson, R. J., & Oakes, T. R. (2006). Motion correction and the use of motion covariates in multiple-subject fMRI analysis. *Human Brain Mapping*, 27(10), 779–788.
- Josephs, O., & Henson, R. N. A. (1999). Event-related functional magnetic resonance imaging: Modelling, inference and optimization. *Philosophical Transactions of the Royal Society of London Series B, Biological sciences*, 354(1387), 1215.
- 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 of London. Series B, Biological sciences*, 362(1481), 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(1), 3–33.
- Meinzer, M., Lahiri, A., Flaisch, T., Hannemann, R., & Eulitz, C. (2009). Opaque for the reader but transparent for the brain: Neural signatures of morphological complexity. *Neuropsychologia*, 47(8–9), 1964–1971.
- Naples, S. E., & Fowler, C. A. (1987). Formal relationships among words and the organization of the mental lexicon. *Journal of Psycholinguistic Research*, 16(3), 257–272.
- Plank, F. (1981). *Morphologische (Ir)Regularitäten*. Tübingen: Gunter Narr.
- Rastle, K., & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes*, 23(7–8), 942–971.
- Rastle, K., Davis, M. H., & New, B. (2004). The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review*, 11(6), 1090–1098.
- Reynolds, C., & Kamphaus, R. (2007). *Test of Irregular Word Reading Efficiency*. Austin, TX: Pro-Ed.
- Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping*, 17(3), 143–155.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niaz, R., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, 23, 208–219.
- Spencer, A. (1991). *Morphological theory: An introduction to word structure in generative grammar*. Oxford: Blackwell.
- Taft, M., & Forster, K. I. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 9, 271–294.
- Torgesen, J. (1999). *Test of word reading efficiency (TOWRE)*. Austin, TX: Pro-Ed.
- Woolrich, M. W. (2008). Robust group analysis using outlier inference. *NeuroImage*, 41(2), 286–301.
- Woolrich, M. W., Behrens, T. E. J., Beckmann, C., Jenkinson, M., & Smith, S. M. (2004). Multilevel linear modelling for FMRI group analysis using Bayesian inference. *NeuroImage*, 21(4), 1732–1747.
- Woolrich, M. W., Jbabdi, S., Patenaude, B., Chappell, M., Makni, S., Behrens, T. E. J., Beckmann, C., Jenkinson, M., & Smith, S. M. (2009). Bayesian analysis of neuroimaging data in FSL. *NeuroImage*, 45(1S1), 173–186.
- Woolrich, M. W., Ripley, B. D., Brady, M., & Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *NeuroImage*, 14(6), 1370–1386.
- Worsley, K. (2001). Statistical analysis of activation images. In: P. Jefferard, P. M. Matthews, & S. M. Smith (Eds.), *Functional MRI: An introduction to methods* (pp. 251–270). Oxford, UK: OUP.