Journal of Memory and Language 63 (2010) 117–130

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Derivational morphology and base morpheme frequency

M.A. Ford [a,*], M.H. Davis [b], W.D. Marslen-Wilson [b]

[a] *Center for Speech and Language, Department of Experimental Psychology, University of Cambridge, Cambridge CB2 3EB, UK*
[b] *MRC Cognition and Brain Sciences Unit, Cambridge, UK*

ARTICLE INFO

Article history:
Received 9 January 2007
revision received 1 December 2008
Available online 24 March 2010

Keywords:
Morphology
Lexical representation
Frequency effects
Lexical decision

ABSTRACT

Morpheme frequency effects for derived words (e.g. an influence of the frequency of the base "dark" on responses to "darkness") have been interpreted as evidence of morphemic representation. However, it has been suggested that most derived words would not show these effects if family size (a type frequency count claimed to reflect semantic relationships between whole forms) were controlled. This study used visual lexical decision experiments with correlational designs to compare the influences of base morpheme frequency and family size on response times to derived words in English and to test for interactions of these variables with suffix productivity. Multiple regression showed that base morpheme frequency and family size were independent predictors of response times to derived words. Base morpheme frequency facilitated responses but only to productively suffixed derived words, whereas family size facilitated responses irrespective of productivity. This suggests that base morpheme frequency effects are independent of morpheme family size, depend on suffix productivity and indicate that productively suffixed words are represented as morphemes.

© 2010 Published by Elsevier Inc.

Introduction

The frequency effect, defined as faster and more accurate recognition of frequently encountered words, provides a useful tool to investigate the representation of language in the mind. It played a crucial role in the development of classic theories of the language system (Broadbent, 1967; Forster, 1976; Morton, 1969). These early theories interpreted the finding that high frequency words are recognised faster than low frequency words in speeded word/non-word classification (lexical decision) (Howes & Solomon, 1951) as evidence that the mental representations of words – a central component of the language processing system – are frequency sensitive.

Early accounts of the mental lexicon treated the lexical representation of each word as a separate atomic entity. However, there are systematic regularities among words that share a common structural element, or morpheme (e.g. */help/* in *helped, helper, helpless*), suggesting that lexical representation is more complex than a frequency sensitive store of discrete word representations. Subsequent models of language comprehension proposed that morphemes are a fundamental unit of lexical representation. They suggest that morphologically complex words are decomposed into their constituent morphemes and that morphologically related words share morphemic representations in the mental lexicon (Giraudo & Grainger, 2000; Kempley & Morton, 1982; Marslen-Wilson, Tyler, Waksler, & Older, 1994; Schreuder & Baayen, 1995; Taft, 1979). This implies that in addition to the robust influence of word frequency on lexical processing, morphemic frequency should also influence lexical processing times. There is now a considerable body of literature suggesting that this is the case (Alegre & Gordon, 1999; Baayen, Dijkstra & Schreuder, 1997; Bradley, 1979; Burani & Caramazza, 1987; Colé, Beauvillain, & Segui, 1989; Meunier & Segui, 1999; Niswander, Pollatsek, & Rayner, 2000; Taft, 1979), supporting a more complex view of the mental lexicon which includes morphemic representations.

* Corresponding author. Fax: +44 (0)1223 766452.
E-mail addresses: mford@csl.psychol.cam.ac.uk, m.ford@uea.ac.uk (M.A. Ford).

0749-596X/$ - see front matter © 2010 Published by Elsevier Inc.
doi:10.1016/j.jml.2009.01.003

Morphology is traditionally divided into inflectional morphology (*cat+s*, *ask+ed*), derivational morphology (*help+less*) and compounding (*black+bird*). Inflectional morphology has been the main focus of psycholinguistic research on the mental representation of morphology. Inflectional endings typically mark syntactic features, such as tense in verbs or number in nouns. Words containing inflectional affixes have forms and meanings that are fully predictable given knowledge of the base and affix. Inflections do not change the semantics or the syntax of the base and show limitless productivity, that is, they are freely attached to novel words to create their inflected forms (e.g. *iPod+s*). This systematic structure makes the representation of inflected words in terms of morphemes a potentially efficient processing strategy. One way this possibility has been investigated is by testing whether lexical decision is influenced by lemma frequency, the summed frequency of the inflectional variants of a base. Facilitatory effects of lemma frequency in lexical decision have been interpreted as evidence for the morphemic representation of inflected words (Alegre & Gordon, 1999; Burani, Salmaso, & Caramazza, 1984; Clahsen, Eisenbeiss, & Sonnenstuhl-Henning, 1997; Colombo & Burani, 2002; Taft, 1979; but see Sereno & Jongman, 1997).

In contrast to inflectional morphology, the relationship in form and meaning between derived words and their base forms is markedly less systematic. The addition of a derivational affix can change both the syntax and semantics of a base (e.g. *govern+ment*), with the resulting forms varying considerably in the predictability of their meaning (e.g. *apart+ment*). Derivational affixes also vary in productivity, for example, the suffix –*ness* (e.g. *cold+ness*) can be attached freely to adjectives to create new nouns but –*th* (e.g. *warm+th*) is no longer used for this function (cf. *blingness*, *blingth*). The unsystematic nature of derivational morphology makes morphemic representation of derived words potentially a less effective strategy for lexical processing than it is for inflected words.

Morpheme frequency effects have been used to test whether derived words activate morphemic representations despite their unsystematic nature. Several different measures of morphemic frequency of derived words have been used in the literature. Some authors have used the sum of the frequencies of some or all morphological variants of a base word, whereas others have simply used base frequency, the whole-form frequency of the base of the derived word (e.g. the influence of the frequency of *help* on responses to *helper*) (Bradley, 1979; Burani & Caramazza, 1987; Colé et al., 1989; Meunier & Segui, 1999).[1] The key claim based on morphemic frequency effects on response times to derived words is the same; however, irrespective of the particular measure as all the different morphemic frequency counts for derived words are highly correlated. Effects of morphemic frequency on lexical decision responses are argued to support models of lexical representation in which base morphemes participate in the representation of derived words, providing a locus at which morphemic frequency effects can accumulate.

Different models of lexical representation have been proposed to account for morphemic frequency effects in lexical decisions to derived words (and effects of morphemic structure in other experimental tasks). In most accounts it is proposed that derived words are represented both as whole-forms and derived words, but at different levels within hierarchical models of lexical representation (Giraudo & Grainger, 2000; Taft, 1994). Most models take account of the variation in consistency that exists in derivational morphology, proposing for example, that only semantically transparent derived words (Marslen-Wilson et al., 1994) activate morphemic representations.

In contrast, Baayen and colleagues propose that only a very limited set of derived words are represented as morphemes (Bertram, Schreuder, & Baayen, 2000). In a series of factorial experiments they investigated response times to semantically transparent Dutch words with inflectional and derivational suffixes. Morphemic frequency effects were found only for words with suffixes which were productive, did not change the meaning of the base or did so only marginally and did not have a homonymic form. For example, facilitatory morphemic frequency effects were found for derived words with the suffix –*heid*, but not for words with the agentive suffix –*er* (Bertram, Schreuder, Baayen, 2000). Both these suffixes are productive and alter the meaning of the base only marginally. However, as in English, the Dutch suffix –*er* is homonymic as it is used to form both agentive noun forms of verbs (*builder*) and the comparative of adjectives (*darker*). As few derivational affixes in English fulfil the criteria proposed by Bertram and colleagues, this predicts that most derived words in English should not show morpheme frequency effects suggesting that they are represented as whole forms and not as morphemes.

Schreuder and Baayen (1997) propose that the inconsistency in the literature on morpheme frequency effects for derived words reflects the fact that morpheme frequency is confounded with morphological family size, the type frequency count of words sharing a morpheme. A word with many morphemic relatives is likely to have a higher morpheme frequency than one with few. In a series of experiments investigating response times to morphologically simple nouns, Schreuder and Baayen (1997) manipulated morphemic frequency and family size. Family size showed a strong facilitatory relationship with response times when cumulative morpheme frequency was controlled, whereas when family size was controlled, effects of base morpheme frequency were not observed (see also Bertram et al., 2000; De Jong, Schreuder, & Baayen, 2000). On the basis of these results they proposed that the majority of derived words would not show base morpheme frequency effects if family size were controlled. They suggested that family size effects reflect feedback from semantics to lexical representations "in the same way as table may become partially activated upon reading chair" (Schreuder & Baayen, 1997, p. 132), rather than shared morphemic representations. As most morphologically related words are also semantically related, morphological family size has a facilitatory relationship with response times.

[1] The affixes in Burani and Caramazza (1987) and Colé et al. (1989) were among the most productive in the respective languages (Italian, French). Bradley (1979) found morphemic frequency effects for words with suffixes classified as productive but not for those classified as non-productive. No details of affix productivity are given in Meunier and Segui (1999).

The current study

The current study investigated the morphological representation of derived words by testing whether base morpheme frequency, family size or both influence responses to derived words in English. In addition, the study also tested whether base morpheme frequency and family size effects depend upon affix productivity, as Bertram and colleagues (Bertram, Schreuder, et al., 2000) found that only productively suffixed words show morpheme frequency effects. In contrast to most previous investigations of morpheme frequency effects, the studies reported here adopted correlational designs. This permitted us to quantify the influence of base morpheme frequency and family size on responses to a large representative sample of derived words without the need to dichotomize these quasi-linear variables. It was predicted that if derived words in English are represented as morphemes, an influence of base morpheme frequency would be found, independently of any influence of family size. In contrast, it was predicted that if Schreuder and Baayen's (1997) findings generalise to derived words in English, the influence of family size should outweigh any influence of base morpheme frequency, indicating that most derived words in English are represented as whole forms, rather than as morphemes. Bertram, Schreuder, et al.'s (2000) results also predict that if any independent influence of base morpheme is found this should be for productively suffixed words only.

Experiment 1

Experiment 1 assessed the influences of base morpheme frequency and family size on lexical decision response times. Orthographic variables and variables measuring the semantic relatedness between derived words and their bases were also fitted in regression models. Multivariate statistics were used for the assessment and treatment of the multicollinearity between predictor variables.

Method

Participants

Participants for all three experiments were selected from the Medical Research Council Cognition and Brain Sciences Unit (MRC-CBU) volunteer panel. The criteria for selection were that participants should be native speakers of British English, aged 17–45 years, without language or hearing problems and with normal or corrected to normal eyesight. Thirty-eight volunteers participated in Experiment 1.

Materials

Word stimuli for all three experiments were selected from the CELEX lexical database (Baayen, Piepenbrock, & Gulikers, 1995). One hundred and eighty semantically transparent derived suffixed words were selected for Experiment 1 (Appendix A). Their properties are summarised in Table 1.

Ratings of the semantic relatedness between the derived words and their bases were obtained either by pretesting or from the CBU Semantic Relatedness Database derived from the original set of ratings collected for the studies reported in Marslen-Wilson et al. (1994). Pre-testing was conducted by asking a separate group of participants from those in the experiment to rate the semantic relatedness of word pairs on a scale of 1 (not at all related) to 9 (highly related). Experimental words comprised approximately 30 percent of the list. The remainder of the list consisted of filler items, which were included to give a broad range of semantic and form overlap to ensure that participants would use the whole range of the scale. The CBU Semantic Relatedness Database consists of the results of many semantic relatedness pre-tests using the same methodology. All derived words in the experiment were rated as highly related in meaning to their bases (mean = 7.8, *SD* = 0.4).

Frequency data were obtained from the CELEX database (Baayen et al., 1995). For each word, the base morpheme frequency, derived word-form frequency and family size were obtained. The CELEX database (Baayen et al., 1995) was also used to calculate the form variables: bigram frequency, trigram frequency and orthographic neighbourhood density, Coltheart's 'N' (Coltheart, Davelaar, Jonasson, & Besner, 1977). Many of the derived words exhibited base allomorphy, that is the base word was not fully embedded in the derived word (*revival/revive*). However, the majority of the allomorphic changes were the common orthographic changes associated with suffixation in English, for example the deletion of final *e* of the base (*revival*). Such frequent and predictable allomorphic changes do not appear to influence the activation of base morphemes (McCormick, Rastle, & Davis, 2007). Only six of the 180 items had base allomorphy involving change to a base consonant (e.g. *prescribe/prescription*), therefore we did not anticipate that allomorphy would interact with the effects of other variables.

Stimuli were divided into words with more or less productive suffixes. Although some affixes are clearly productive or non-productive, many affixes are marginally productive, that is, they are occasionally used to form new words. Therefore, ideally we would have used a continuous measure as we have done with our other predictors. However, the currently available measures are often contradictory, with one measure indicating high productivity for an affix and another indicating low productivity (Plag, 2004). For this reason, we chose to dichotomize pro-

Table 1
Experiment 1, stimuli characteristics.

	WF	BF	FS	Len	Syll	'N'	BG	TG
Mean	6.9	92.5	4.0	8.5	2.9	0.6	31,611	4542
SD	7.4	107.6	2.5	1.7	0.8	1.3	8744	2535
Max	67	431	11	13	5	8.0	73,773	24,520
Min	0	1	1	5	2	0.0	13,981	662

WF, word-form frequency of derived word; BF, base morpheme frequency; FS, family size; Len, number of letters; Syll, number of syllables; 'N', neighbourhood density; BG, bigram frequency; TG, trigram frequency. Data from CELEX. Frequencies are per million.

ductivity and used a number of sources and methods to aid the reliability of the categorisation as the different ways of assessing productivity we used are likely to "highlight different aspects of productivity" (Plag, 2004, p. 16). The measures we used to assess productivity included the number of hapax legommena of an affix, the type and token frequency of an affix and dictionary citation dates for neologisms (Baayen & Lieber, 1991; Baayen et al., 1995; Bauer, 1983; Marchand, 1969; Oxford English Dictionary on CD-ROM, 1992). As dichotomisation requires a somewhat arbitrary division into more and less productive suffixes, three different categorisations were used. The first productivity split contrasts words with highly productive suffixes with all others. The second split contrasted words with moderately to highly productive suffixes with all others. The third split contrasted words with marginally to highly productive suffixes with all others. Details of the productivity splits are given in Appendix B.

The experiment included 180 real word filler words: 90 morphologically simple words matched in length and frequency to the derived forms and 90 morphologically simple words matched in length and frequency to the bases of the derived words. The real word fillers were included to reduce the proportion of derived words in the experimental lists.

A set of 540 non-words was created. Two sets of 180 non-word foils were pair-wise matched to the derived words and their bases in length, number of syllables and neighbourhood density. Each derived non-word was created by changing one or two letters of the base of a derived word matched to a test item, in order to preserve the suffix, thus creating foils with non-word bases and real suffixes (e.g. *brishly*). Morphologically simple non-words were created in a similar manner. A further 180 non-words were created to match the derived words and their bases as a group, in length, syllables, bigram frequency and neighbourhood density.

The test items were divided into two experimental versions to avoid participants seeing large numbers of items with the same suffix. Derived words and their bases were rotated across the versions so that each version contained 90 derived words and 90 bases. The pair-wise matched non-word fillers were also rotated so that each version included 90 non-word derived items and 90 non-word bases. This ensured that suffixation could not be used to predict lexicality. All other fillers were the same in both versions. Both versions began with a short practice session to accustom participants to the task. Each version was divided into six experimental blocks to allow participants to have frequent breaks. Each block began with six warm-up trials. In total, both versions had 776 items: 20 practice items, 36 warm-up items, 90 base items, 90 derived items, 180 real word fillers and 360 non-word foils.

Procedure

A single word visual lexical decision task was used. Items were presented to participants on a computer monitor using DMDX experimental software (Forster & Forster, 2003). Each item was preceded for 250 ms with a "+" fixation point, followed by the item presented for 500 ms in upper case 14 point Arial font. Participants were instructed to make a lexical decision to each item as quickly and accurately as possible by pressing one of two keys on a button box. The 'yes' response key was always pressed with the dominant hand and the 'no' response with the non-dominant hand. Participants had 2000 ms in which to respond before the programme timed out and moved onto the next item. There was a minimum inter-trial interval of 2000 ms. Each participant saw a different pseudo-randomised order of the list.[2]

Results of Experiment 1

The criteria for removal of participants for all three experiments was an overall error rate exceeding 15% or an error rate for either words or non-words alone exceeding 20%, possibly indicating a response bias. The criterion for removal of items was an error rate exceeding 30%. Using these criteria data from four participants and three items (*allegation, decoration, lender*) were removed from the analyses. This left a total of 34 participants and 177 items. All lexical decision errors, i.e. false rejections of word targets (5.3% of data) were removed from the data. Mean item and participant response times were log-transformed and entered into analyses. Only analyses of response times will be reported here as analyses of error data in all three experiments showed no significant effects of interest to the current study (e.g. frequency effects). A summary of the correlations between predictor variables and average response times is presented in Table 2. Only analyses of response times will be reported here as analyses of error data in all three experiments showed no significant effects of interest to the current study (e.g. frequency effects).

Regression analyses

The main regression analyses were by items analyses, in which the dependent variable was the response times to words averaged across participants. In order to reduce possible problems with collinearity in regressions including both the word-form frequency of the derived words and their base morpheme frequency, a regression was carried out with base frequency as the dependent variable and derived word frequency as the independent variable. The standardised residual of this regression was saved as a measure of base morpheme frequency with derived word-form frequency partialled out. Hierarchical regression models were used in which residualized base morpheme frequency was fitted in the second block and all other predictors fitted in the first block. In addition, regressions of results by participants were carried out. In these analyses separate standard regressions were conducted on each participant's response time data, fitting the same predictor variables used regressions by items. Participants'

[2] A reviewer was concerned that the experiment was rather long and that despite the participants having frequent breaks fatigue might be a problem. However, one sample *t*-tests on by-participant Pearson and Spearman correlations between response times and list order were not significant ($ts < 1$). This suggests that using separate randomisations for each participant and giving frequent breaks successfully prevented presentation order from influencing response times.

Table 2
Experiment 1, Pearson correlations.

	WF	BF	FS	SR	FPCA1	FPCA2
RT	−0.253[a]	−0.349[a]	−0.311[a]	−0.084	0.420[a]	−0.111
WF		0.261[a]	0.118	−0.011	0.076	−0.185[b]
BF			0.386[a]	−0.142[c]	−0.149[b]	0.056
FS				−0.092	−0.162[b]	−0.033
SR					−0.067	0.001
FPCA1						0.000

[a] Correlation is significant at the 0.01 level (1-tailed).
[b] Correlation is significant at the 0.05 level (1-tailed).
[c] Correlation is significant at the 0.1 level. WF, derived word-form frequency; BF, base morpheme frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2.

semi-partial correlation co-efficients for the predictors of interest were then compared using ANOVAs to assess whether effects were consistent across participants (Lorch & Myers, 1990).

The predictors entered in the items analyses were the word-form frequency of the derived word, residualized base morpheme frequency, family size, the suffix productivity dummy variable, the semantic relatedness of the derived words and their base words and form variables. A principal component analysis was previously carried out on the form variables to reduce the number of predictors (cf. Hauk, Davis, Ford, Pulvermüller, & Marslen-Wilson, 2006). This resulted in two components which were used in these analyses. The first component was bipolar with a high positive correlation with the number of letters and a high negative correlation with neighbourhood density. The second component had a strong positive correlation with bigram and trigram frequency. Data for other predictor variables were converted to *Z*-scores in all three experiments to further reduce any multicollinearity between variables to acceptable levels (tolerance > 0.25, condition index < 15) (Stevens, 1999). In the analyses for Experiment 1, the maximum condition index was 5.6 and the maximum variable inflation factor was 1.4. All variables except semantic relatedness and form PCA variable 2 were significantly correlated with response times – note that suffix productivity is not included in this preliminary correlation analysis as it is a binary variable.

Turning to the main regression analysis, this was based on the second productivity split, which had the most balanced sets of words with more and less productive suffixes (85 words with moderately to highly productive suffixes versus 92 words with marginally or non-productive suffixes). *R* for regression (measuring the overall significance of the regression) was significantly different from zero ($F(8, 168) = 16.4$, $p < .001$, total $R^2 = 0.438$) (see Table 3). In the first block derived word-form frequency, family size and the first and second form PCA variables were significant facilitatory predictors of response times ($\beta = -0.353$, $p < .001$; $\beta = -0.346$, $p < .001$; $\beta = -0.192$ $p < .01$; $\beta = -0.119$, $p < .05$), with high values on these variables associated with faster response times. The dummy variable coding productivity was also a significant predictor ($\beta = -0.127$, $p < .05$) indicating that words with more productive suffixes were responded to faster than those with less productive suffixes. Semantic relatedness was not a significant predictor of response times ($\beta = -0.074$, $p > .1$). Subsequent fitting of the residualized base frequency did not significantly add to the explained variance of the model (change in $R^2 = 0.002$, $p > .1$), consistent with the predictions of Baayen and colleagues that few derived words should show effects of morphemic frequency (Bertram, Schreuder, et al., 2000; Schreuder & Baayen, 1997). The results of regression analyses for the other productivity splits confirmed this pattern of results. As there was no difference between the different productivity divisions used in this or subsequent analyses, from now on we focus on the results of analyses using the second categorisation of more and less productive suffixes (e.g. moderately to highly productive suffixes versus marginally or non-productive suffixes).

Given the hypotheses motivating this study, further items regression analyses were conducted to investigate whether productivity interacted with base morpheme frequency, family size or semantic relatedness. Analyses were conducted as above but subsequent interaction terms were fitted. For base morpheme frequency, the interaction term for productivity added significantly to the explanatory power of the equation (change in $R^2 = 0.014$, $p < .05$). The interaction terms for family size or semantic relatedness were not significant ($ps > .1$). A similar pattern of results was also found with the other two productivity splits. Additionally the interaction of base morpheme frequency and productivity was consistent across participants. Repeated measures analyses of variance were conducted on participants' semi-partial correlation co-efficients for the variables of interest (e.g. base morpheme frequency, family

Table 3
Experiment 1, hierarchical regression, Block 1.

	Co-efficients			*t*	Sig.	Tol	VIF
	R	Std. Err.	*β*				
Version	−0.020	0.006	−0.209	−3.6	0.000	0.988	1.012
FPCA1	0.017	0.003	0.353	5.8	0.000	0.908	1.102
FPCA2	−0.006	0.003	−0.119	−2.0	0.043	0.987	1.013
WF	−0.017	0.003	−0.346	−5.8	0.000	0.946	1.057
FS	−0.009	0.003	−0.192	−3.2	0.002	0.924	1.082
SR	−0.004	0.003	−0.074	−1.3	0.209	0.973	1.028
PROD2	−0.012	0.006	−0.127	−2.1	0.038	0.907	1.102

R, regression co-efficient; Std. Err., standard error of the regression co-efficient; *β*, standardised regression co-efficient; *t*, *t* statistic; Sig., significance of *t* statistic; Tol, tolerance; VIF, variable inflation factor; WF, derived word-form frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2; PROD2, productivity dummy variable 2.

size, semantic relatedness and derived word frequency). There was a main effect of productivity for base morpheme frequency ($F1(1, 32) = 5.2$, $p < .05$) but not the other three variables (all $ps > .1$).

As Baayen, Feldman, and Schreuder (2006) report non-linear effects of family size we also carried out an analysis where we fitted a quadratic term for family size in the second block. Fitting quadratic family size did not add significantly to the explained variance of the model (change in $R^2 = 0.004$, $p > .1$). However, the quadratic effect reported by Baayen and colleagues was due to a facilitatory effect at the lower range of family size and an inhibitory effect at in the higher range of family size. It is possible that the range of family size in Experiment 1 was not sufficiently large for the inhibitory effects for large family sizes to be detected.

Post hoc tests were conducted to test whether the results of Experiment 1 were limited to the words which had the more productive and non-homonymic suffixes, as would be predicted by Bertram and colleagues (Bertram, Schreuder, et al., 2000). An interaction term was created using a dummy variable coding words with more productive and non-homonymic suffixes. We classified a suffix as homonymic if it had at least one other clearly distinct meaning or usage. Thus, *–er* was defined as homonymic due to the existence of the competing comparative suffix *–er*. The suffix *–less* was classified as being non-homonymic because despite attaching to both nouns and verbs, it always creates adjectives and there is substantial overlap in meaning between the form that attaches to nouns (*faultless*) and the form that attaches to verbs (e.g. *tireless*). After fitting the original variables and the binary homonymy variable in the first block of a hierarchical model and residualized base frequency in the second block, fitting the interaction term in a third block did not add significantly to the explained variance (change in $R^2 = 0.007$, $p > .1$).

Discussion

Experiment 1 found that family size and base morpheme frequency were significant independent facilitatory predictors of response times to derived suffixed words and that affix productivity is an important determinant of morphemic frequency effects. The interaction of base morpheme frequency and affix productivity was significant across both items and participants, irrespective of which classification of productivity was used. However, the amount of explained variance added to the model by fitting the base morpheme frequency by productivity interaction term was small suggesting that base morpheme frequency has only a limited influence on response times to derived words over and above that of the frequency of suffixed word itself and its morphological family size.

In addition, the results of Experiment 1 suggest that affix productivity may be more important than affix homonymy in determining whether a derived word will show an effect of base morpheme frequency. However, this conclusion must remain tentative as Experiment 1 was not designed to explicitly examine effects of homonymy on the processing of derived words and does not address issues such as how the relative frequencies of particular affixes may influence competition between them.

Separate items analyses of responses to words with more and less productive suffixes could not be conducted for this set of materials. The use of three different categorisations of productivity at the time of selecting these items precluded using a fully balanced set of words with more and less productive suffixes as would be required for a satisfactory analysis in which these two classes are analysed separately.

Experiment 2

A second experiment was conducted using larger, balanced sets of words with more and less productive suffixes to investigate responses to these classes of words separately. The interaction between base morpheme frequency and affix productivity found in Experiment 1 predicts that base morpheme frequency should only facilitate responses to words with more productive suffixes. Family size, in contrast, viewed as a non-morphological variable, should show similar effects for both word types.

Method

Participants

Forty-eight volunteers from the MRC-CBU volunteer panel participated.

Materials

One hundred and eight words with more productive and 108 less productive suffixes were selected (Appendix C). Productivity was categorised according to the second definition of Experiment 1, so that the more productive set included words with derivational suffixes that are only moderately productive in current English, e.g. *–ment*. The two sets of suffixed words were matched on semantic relatedness, frequency and form variables and are summarised in Table 4.

The test items were divided into two versions of the experiment to avoid participants seeing large numbers of items with the same suffix. Each version had 108 tests words (54 with more productive and 54 with less productive suffixes). In addition, each list contained two sets of 108 monomorphemic fillers matched in length, syllables and frequency to the derived words and their bases, respectively. A set of 324 non-words, matched to the test items and real word fillers using the same criteria as in Experiment 1, was also included.

Procedure

The procedures followed were the same as in Experiment 1.

Results

Using the criteria for item and participant removal described above, the data from seven participants and four items were excluded (*dissenter, dampish, deportation, lender*). All seven participants were from Version 1. This left

Table 4
Experiment 2, stimuli characteristics.

		WF	BF	FS	Ln	Syll	'N'	BG	TG	SR
More productive	*M*	7.9	75.1	4.0	8.5	2.7	1.1	31,652	4922	8.0
	Std	8.8	76.2	2.8	1.8	0.8	1.8	7620	2097	0.5
	Min	0	2	1	5.0	0.0	0.0	14,022	662	6.7
	Max	67	296	16	13.0	5.0	9.0	56,521	11,347	8.9
Less productive	*M*	8.0	62.5	4.4	8.5	3.1	0.4	31,775	4135	7.9
	Std	6.8	79.3	3.5	1.5	0.7	0.8	9067	2752	0.4
	Min	0	1	1	5.0	1.0	0.0	13,981	1090	7.1
	Max	33	357	28	13.0	5.0	4.0	73,773	24,521	8.8

WF, word-form frequency of derived word; BF, base morpheme frequency; FS, family size; Len, number of letters; Syll, number of syllables; 'N', neighbourhood density; BG, bigram frequency; TG, trigram frequency. Data from CELEX. Frequencies are per million.

a total of 41 participants (21 in Version 1 and 20 in Version 2) and 212 items. All lexical decision errors, i.e. false rejections of word targets (3.6% of data) were removed from the data. Mean item and participant response times were log-transformed and entered into analyses.

Analyses of variance

By item and by participant analyses of variance showed that responses to words with more productive suffixes were significantly faster than responses to words with less productive suffixes ($F1(1,39) = 39.5$, $p < .001$, $F2(1,210) = 6.8$, $p < .02$) (Fig. 1). There was no significant difference in error rates between words with more and less productive suffixes, however ($F1(1,39) = 2.2$, $p > .1$, $F2(1,210) = 1.7$, $p > .1$).

Regression analyses for derived words with more productive suffixes

The same hierarchical procedure was used to analyse the items data as in Experiment 1, with the residualized base morpheme frequency entered in a second block after fitting all other variables simultaneously the first block. The predictors entered in the first block were family size, derived word frequency, the semantic relatedness of the derived words and their bases and form variables. A summary of the correlations between variables and response times is presented in Table 5. In the analyses in Experiment 2, the maximum condition index was 6.3 and the maximum variable inflation factor was 1.3.

R for regression was significantly different from zero ($F(7,96) = 10.4$, $p < .001$, total $R^2 = 0.431$) (Table 6). Derived word frequency and form PCA variable 1 were significant predictors of response times ($\beta = -0.477$, $p < .001$; $\beta = -0.355$, $p < .001$), but family size was only a marginal predictor of response times ($\beta = -0.147$, $p = .08$). Neither semantic relatedness nor form PCA variable 2 were significant predictors of response times ($\beta = -0.091$, $p > .1$; $\beta = 0.011$, $p > .1$). Consistent with Experiment 1, subsequently fitting base morpheme frequency added significantly to the explained variance of the model, although this effect was marginally significant (change in $R^2 = 0.016$, $p = .09$). Two additional analyses were carried out. In the first, a quadratic term for family size was fitted in the second block of a hierarchical analysis. This did not significantly add to the explained variance of the model (change in $R^2 = 0.001$, $p > .1$). The second additional analysis investigated whether the marginal effect of base fre-

Table 5
Experiment 2, correlations for productively suffixed words.

	WF	BF	FS	SR	FPCA1	FPCA2
RT	−0.312[a]	−0.348[a]	−0.302[a]	−0.023	0.356[a]	−0.108
WF		0.325[a]	0.109	−0.160	0.202[b]	0.193[c]
BF			0.292[a]	−0.373[a]	−0.153	0.189[c]
FS				−0.042	−0.235[b]	−0.073
SR					−0.046	−0.214[b]
FPCA2						0.000

[a] Correlation is significant at the 0.01 level (1-tailed).
[b] Correlation is significant at the 0.05 level (1-tailed).
[c] Correlation is significant at the 0.1 level. WF, derived word-form frequency; BF, base morpheme frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2.

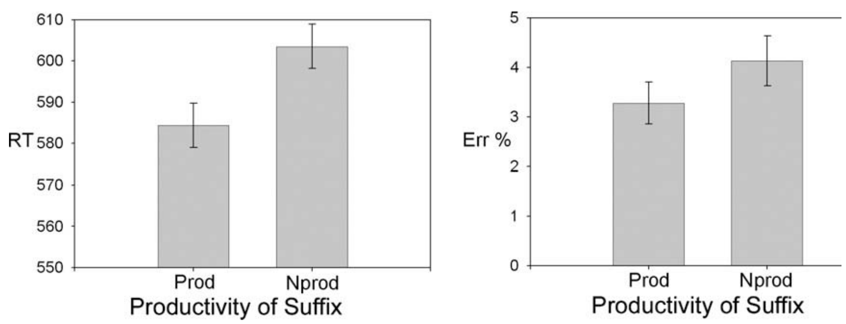

Fig. 1. Experiment 2, mean reaction times (RT) and percentage error (Err%) for words with more productive (MoreProd) and less productive (LessProd) suffixes.

Table 6
Experiment 2, Block 1, words with more productive suffixes.

	Co-efficients			*t*	Sig.	Tol	VIF
	R	Std. Err.	*β*				
Version	−0.047	0.013	−0.272	−3.5	0.001	0.983	1.018
WF	−0.041	0.007	−0.477	−5.6	0.000	0.820	1.219
FPCA1	0.031	0.007	0.355	4.4	0.000	0.915	1.093
FPCA2	0.001	0.007	0.011	0.1	0.902	0.807	1.239
FS	−0.013	0.007	−0.147	−1.8	0.075	0.909	1.100
SR	−0.008	0.007	−0.091	−1.1	0.259	0.938	1.066

R, regression co-efficient; Std. Err., standard error of the regression co-efficient; *β*, standardised regression co-efficient; *t*, *t* statistic; Sig., significance of t statistic; Tol, tolerance; VIF, variable inflation factor. WF, derived word-form frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2.

quency was due to the inclusion of words with non-homonymic suffixes. After fitting the original variables and a binary homonymy variable in the first block of a hierarchical model and residualized base frequency in the second block, subsequently fitting the interaction term in a third block did not add significantly to the explained variance (change in $R^2 = 0.005$, $p > .1$).

Regression models in which only family size or base morpheme frequency was fitted suggested that the weakness of the base morpheme frequency effect may reflect joint variance. In a hierarchical regression model with a first block as above but with family size excluded, fitting base morpheme frequency significantly added to the explained variance of the model (change in $R^2 = 0.026$, $p < .05$). In a hierarchical regression with a first block as above but family size not included, the effect of fitting family size in a second block was still marginal (change in $R^2 = 0.020$, $p = .08$). Two further hierarchical regressions were conducted. In the first model, all predictors other than base frequency and family size were fitted in the first block, family size in the second block and base morpheme frequency in the third block. In the second hierarchical regression, the order of fitting base morpheme frequency and family size was reversed. When base morpheme frequency was fitted before family size, it added significantly to the explained variance (change in $R^2 = 0.026$, $p < .05$) but family size did not add significantly to the explained variance (change in $R^2 = 0.011$, $p > .1$). When family size was fitted before base morpheme frequency it added to the explained variance marginally (change in $R^2 = 0.020$, $p = .08$) but base morpheme frequency still added to the explained variance marginally (change in $R^2 = 0.018$, $p = .09$). The hierarchical regressions thus provide some support for base morpheme frequency as a significant predictor of response times to derived words with more productive suffixes when effects of family size are partialled out.

As Experiment 2 used two different experimental lists (versions), the regression co-efficient (constant) from a one-way analysis of variance was used as a measure of the consistency of the semi-partial correlation co-efficients across participants (Matthews, Altman, Campbell, & Royston, 1990). If the constant is significant, it demonstrates that the mean semi-partial correlation co-efficient is reliably different from zero. All four variables of interest showed consistent facilitatory effects across participants (base morpheme frequency, $F(1, 39) = 4.7$, $p < .05$; family size, $F(1, 39) = 10.0$, $p < .01$; semantic relatedness $F(1, 39) = 8.2$, $p < .01$; derived word frequency $F(1, 39) = 52.8$, $p < .001$).

Regression analyses for derived words with less productive suffixes

A summary of the correlations between variables and response times is presented in Table 7.

R for regression was significantly different from zero ($F(7, 100) = 5.9$, $p < .001$, total $R^2 = 0.292$) (Table 8). Derived word frequency and form PCA variable 1 were significant facilitatory predictors of response times ($\beta = -0.358$, $p < .001$; $\beta = -0.305$, $p < .001$). The other variables in block 1, family size, semantic relatedness and form PCA variable 2 did not influence response times significantly (Table 8). Fitting residualized base frequency in the second block did not significantly add to the explained variance of the model (change in $R^2 = 0.002$, $p > .1$). In order to check for effects of joint variance, hierarchical regression models were carried with the first block as in the main analysis above but with family size excluded. Subsequently fitting base morpheme frequency did not add significantly to the explained variance of the model (change in $R^2 = 0.005$, $p > .1$) nor did subsequently fitting family size (change in $R^2 = 0.008$, $p > .1$). However, fitting a quadratic term for family size added significantly to the explained variance (change in $R^2 = 0.031$, $p < .05$).

By participants analyses using the method described above confirmed the findings of the items analyses showing that effects of base morpheme frequency, family size and semantic relatedness were not consistent across par-

Table 7
Experiment 2, correlations for less productively suffixed words.

	WF	BF	FS	SR	FPCA1	FPCA2
RT	−0.303[a]	−0.185[c]	−0.235[b]	−0.134	0.317[a]	−0.043
WF		0.231[b]	0.129	0.093	−0.027	−0.128
BF			0.395[a]	0.101	0.014	0.043
FS				0.021	−0.224[b]	0.017
SR					−0.093	0.073
FPCA2						0.000

[a] Correlation is significant at the 0.01 level (1-tailed).
[b] Correlation is significant at the 0.05 level (1-tailed).
[c] Correlation is significant at the 0.1 level. WF, derived word-form frequency; BF, base morpheme frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2.

Table 8
Experiment 2, hierarchical regression, Block 1, words with less productive suffixes.

	Co-efficients			*t*	Sig.	Tol	VIF
	R	Std. Err.	*β*				
Version	−0.044	0.015	−0.257	−3.0	0.003	0.958	1.044
WF	−0.031	0.007	−0.358	−4.2	0.000	0.963	1.038
FPCA1	0.026	0.007	0.305	3.5	0.001	0.927	1.079
FPCA2	−0.004	0.007	−0.049	−0.6	0.562	0.992	1.008
FS	−0.008	0.007	−0.090	−1.0	0.298	0.952	1.051
SR	−0.002	0.007	−0.022	−0.3	0.801	0.945	1.059

R, regression co-efficient; Std. Err., standard error of the regression co-efficient; *β*, standardised regression co-efficient; *t*, *t* statistic; Sig., significance of *t* statistic; Tol, tolerance; VIF, variable inflation factor. WF, derived word-form frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2.

ticipants ($F(1, 39) = < 1$; $F(1, 39) = 2.1$, $p > .1$; $F(1, 39) < 1$, respectively). Derived word frequency, however, did show a consistent facilitatory effect across participants ($F(1, 39) = 34.8$, $p < .001$).

Discussion

Experiment 2 investigated base morpheme frequency effects for large, balanced sets of words with more and less productive suffixes. Base morpheme frequency effects were only a significant predictor of lexical decision response times for the words with more productive suffixes. Family size effects were a significant predictor of responses to words with both more productive suffixes (linear) and less productive suffixes (quadratic). The effect of both base morpheme frequency and linear effects of family size were noticeably small. As in Experiment 1, the results of Experiment 2 suggest that although both type and token morphemic frequency do influence responses to derived words, this influence is small when compared to that of the frequency of the derived word itself.

Experiment 3

In both Experiments 1 and 2 no significant effects of base morpheme frequency were found with words for words with suffixes of limited productivity. It is possible, however, that an influence of base morpheme frequency on response times to words with less productive suffixes does exist but is relatively weak and was therefore not detected by Experiments 1 and 2. In Experiment 3, a large set of derived words with suffixes of limited productivity were selected to maximise the possibility of finding a base morpheme frequency effect, if one exists.

Method

Participants

Twenty-seven volunteers from the MRC-CBU volunteer panel participated.

Materials

One hundred and nineteen semantically transparent words with suffixes of limited productivity were selected to maximise the ratio of base morpheme frequency to derived whole-word frequency, in order to increase the likelihood of detecting any influence of base morpheme frequency (Appendix D). The mean ratio of the base morpheme frequency to derived word frequency is 76:1, although the median of 33:1, is a better guide to this, as the mean is somewhat inflated by a few items with very high ratios. Although this ratio might be a useful predictor in a regression model, we did not include it in the regression models as it naturally correlates with both the word-form frequency of derived word and residualized base frequency and thus would increase the likelihood of multicollinearity. Stimuli properties are summarised in Table 9.

One hundred and eighty-one real word filler items were also included, 91 morphologically simple real filler items matched on frequency and form variables to the derived items and 90 morphologically simple items matched on frequency and form variables to the bases of the derived items. Three hundred non-word fillers were also included in the experiment and were matched to the properties of the real words in the same manner as in Experiments 1 and 2. In addition, there was a practice block of 24 items and 36 warm-up items, split equally into words and non-words.

Table 9
Experiment 3, stimuli characteristics.

	WF	BF	FS	Ln	Syll	'N'	BG	TG	SR
M	3.7	65.4	8.1	8.2	2.8	0.6	32947	4166	7.6
Std	5.4	70.4	8.0	1.7	0.7	1.1	9021	2574	0.4
Min	0	1	1.0	5.0	2.0	0.0	17853	995	6.7
Max	33	344	54.0	12.0	5.0	6.0	61603	19956	8.8

WF, word-form frequency of derived word; BF, base morpheme frequency; FS, family size; Len, number of letters; Syll, number of syllables; 'N', neighbourhood density; BG, bigram frequency; TG, trigram frequency. Data from CELEX. Frequencies are per million.

Procedure

The procedures followed were the same as in Experiment 1.

Results and discussion of Experiment 3

Using the criteria for item and participant removal described above, the data from two items were removed (*olden, treasonous*). This left a total of 27 participants and 117 items. All lexical decision errors, i.e. false rejections of word targets (4.7% of data) were removed from the data. Mean item and participant response times were log-transformed and entered into analyses.

Regression analyses

The method of analysis was as above, with hierarchical models fitting residualized base morpheme frequency in a second block after an initial simultaneous fitting of all other variables. The predictors entered were family size, derived word frequency, the semantic relatedness of the derived words and their bases, form variables and residualized base frequency. A summary of the correlations between variables and response times is presented in Table 10. In the analyses for Experiment 3, the maximum condition index was 2.4 and the maximum variable inflation factor was 1.5.

R for regression was significantly different from zero ($F(6, 110) = 10.4$, $p < .001$, total $R^2 = 0.361$) (Table 11). Derived word frequency was a significant facilitatory predictor of response times ($\beta = -0.421$, $p < .001$). Consistent with Experiment 1 and the claims of Baayen and colleagues (Bertram, Baayen, & Schreuder, 2000; Schreuder & Baayen, 1997) family size was also a significant facilitatory predictor of response times ($\beta = -0.276$, $p < .01$). Neither form variables nor semantic relatedness were significant predictors of response times (*ps* > .1). Fitting base morpheme frequency in a second block did not add to the explained variance (change in $R^2 = 0.007$, $p > .1$).

Two further items analyses were carried out. The first tested whether there was any non-linear effect of family size by fitting a quadratic term in a hierarchical regression model. The first block was as above and the quadratic term was fitted in a second block. Quadratic family size did not add to the explained variance of the model (change in $R^2 = 0.008$, $p > .1$). According to Baayen et al. (2006) the quadratic effect of family size is due to a facilitatory effect for words with in the low range of family sizes and an inhibitory effect for words with large family sizes. Experiment 3 had the highest mean family size of this study (8.1), yet surprisingly no quadratic effects of family size were found. The second analysis investigated whether the lack of an effect of base morpheme frequency was a result of the inclusion of derived words which do not fully embed their base morpheme (e.g. *dosage/dose*). When only those derived words which fully embed their base morpheme ($N = 90$) are included in the model, fitting base morpheme frequency still does not add to the variance in the model (change in $R^2 = 0.00$).

One-sample *t*-tests (against a test value of zero) on participant's semi-partial correlation co-efficients showed that there was a marginally consistent inhibitory effect of base morpheme frequency ($t(26) = 1.9$, $p = .06$). However, the confidence intervals for this analysis contain zero, suggesting that this is not a reliable result. Effects of semantic relatedness and the second form PCA variable were not consistent across participants (both, $t(26) < 1$). The first form PCA variable also showed a marginally consistent inhibitory effect ($t(26) = 2.0$, $p = .06$), but again, the confidence intervals suggested that this effect was not reliable. Both derived word frequency and family size, however, showed consistent facilitatory effects across participants ($t(26) = -6.8$, $p < .001$; $t(26) = -5.0$, $p < .001$, respectively).

Table 10
Experiment 3, Pearson correlations.

	WF	BF	FS	SR	FPCA1	FPCA2
RT	−0.346[a]	−0.380[a]	−0.395[a]	−0.141	0.273[a]	0.016
WF		0.470[a]	0.024	0.433[a]	0.051	−0.132
BF			0.541[a]	0.068	−0.217[b]	−0.085
FS				−0.106	−0.323[a]	0.071
SR					0.025	−0.017
FPCA2						0.000

WF, derived word-form frequency; BF, base morpheme frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2.
[a] Correlation is significant at the 0.01 level (1-tailed).
[b] Correlation is significant at the 0.05 level (1-tailed).

Table 11
Experiment 3, hierarchical regression, Block 1.

	Co-efficients			*t*	Sig.	Tol	VIF
	R	Std. Err.	*β*				
WF	−0.032	0.007	−0.421	−4.9	0.000	0.791	1.264
FPCA1	0.008	0.006	0.108	1.3	0.188	0.880	1.137
FPCA2	0.000	0.006	0.000	0.0	0.996	0.984	1.016
FS	−0.021	0.006	−0.276	−3.3	0.001	0.839	1.192
SR	−0.002	0.006	−0.023	−0.3	0.780	0.837	1.194

R, regression co-efficient; Std. Err., standard error of the regression co-efficient; *β*, standardised regression co-efficient; *t*, *t* statistic; Sig., significance of *t* statistic; Tol, tolerance; VIF, variable inflation factor. WF, derived word-form frequency; FS, family size; SR, semantic relatedness score; FPCA1, form PCA variable 1; FPCA2, form PCA variable 2.

General discussion

The current study investigated lexical representation using a correlational design to test the influence of multiple frequency measures on responses to a large, representative sample of derived words in English. Morphology provides regularities in the relationship between the form and meaning of words, from which the language system may be able to profit if morphologically complex words are represented as morphemes. Derivational morphology, in contrast to inflectional morphology, varies in the consistency of the relationship between form and meaning. Therefore it may be efficient to represent as morphemes only those derived words showing a consistent relationship between form and meaning. This predicts that some, but not all, derived words will show morpheme frequency effects.

The variation in the consistency of the mapping of form and meaning for derived words is in part associated with the productivity of the affixes. More productive affixes typically show greater consistency in their form-meaning relationship than less productive affixes. For example, the productive suffix –*ness* has a single, mainly syntactic function, creating a noun out of an adjective and has only one semantically opaque form (*business*). In contrast, the less productive affix –*age* attaches to both nouns (*vicarage*) and verbs (*breakage*), adds to the meaning of the base in an inconsistent fashion (cf. *vicarage, breakage*) and has a number of opaque forms (*footage*, *bandage*). Morphemic representation would appear to be more likely for words with highly productive affixes than those that exhibit little or no productive use.

In Experiment 1 we found an interaction between the effects of base morpheme frequency and suffix productivity. This interaction was significant irrespective of where the division was made between more and less productive suffixes. Experiment 2 compared responses to balanced sets of derived words that varied in suffix productivity. In this experiment base morpheme frequency effects were found only for the words with more productive suffixes. In Experiment 3 we selected derived words with less productive suffixes that had very high ratios of base to derived word-form frequency to maximise the possibility of detecting base frequency effects. However, although there was an effect of family size in this experiment, there was no significant effect of base morpheme frequency.

The finding of a facilitatory base morpheme frequency effect is consistent with previous reports of morphemic frequency effects for derived words (Bradley, 1979; Burani & Caramazza, 1987; Colé et al., 1989; Meunier & Segui, 1999). The interaction between base morpheme frequency and productivity is consistent with the claim of Bertram and colleagues that productivity is an important determinant of the representation of morphologically complex words (Bertram, Schreuder, et al., 2000). Post hoc analyses of Experiment 1 and Experiment 2 suggested, that affix homonymy may not be as important a determinant of the representation of derived words as previously proposed (Bertram, Schreuder, et al., 2000). However, the experiments in this study were not designed to specifically examine homonymy, therefore this conclusion must remain tentative.

Schreuder and Baayen (1997) suggest that previous findings of base morpheme frequency effects do not reflect morphemic representation of derived words. They propose that base morpheme frequency is typically confounded with morphological family size and that family size effects are primarily semantic (Bertram, Baayen, et al., 2000; De Jong et al., 2000; Schreuder & Baayen, 1997). In this study, however, derived words showed independent facilitatory effects of both family size and base morpheme frequency after these variables were transformed to reduce multicollinearity. In addition, we found that the base morpheme frequency of words with more productive suffixes added to the variance explained by hierarchical regression models after the whole-form frequency of the derived word and its morphological family size had already been fitted. This suggests that derived words with more productive suffixes are represented in terms of their constituent morphemes.

The results of this study add to the growing body of evidence indicating that morphological family size is an important influence on lexical processing. Both Experiments 1 and 3 found robust linear effects of family size, but no non-linear effects. In Experiment 2, there was a small, linear effect of family size for the words with more productive suffixes. No linear effect of family size was found for words with less productive suffixes, but there was a significant quadratic effect.

Our failure to replicate in Experiment 3 the quadratic effect of family size found words with less productive suffixes in Experiment 2 is surprising as the findings in Baayen et al. (2006) suggest that the greater number of high family size items in Experiment 3 should have led to stronger quadratic effects of family size. Nevertheless, the significant quadratic effects of family size in Experiment 2 suggest that studies investigating morphemic family size should test for non-linear effects but also that further study is needed to investigate the nature of these non-linear effects.

The absence of an interaction between family size and productivity supports the claim that family size effects do not reflect morphological processes (Schreuder & Baayen, 1997). Previous studies reporting morphemic frequency effects did not control for family size (Bradley, 1979; Burani & Caramazza, 1987; Colé et al., 1989; Meunier & Segui, 1999). The results of the current study suggest that both base morpheme frequency and morphological family size influence response times to derived words. The base morpheme frequency effect suggests that morphologically related words share morphemic representations, whereas the family size effect may reflect the shared semantic representations for morphologically related words (Schreuder & Baayen, 1997).

Words with more productive suffixes showed both a base morpheme frequency effect and a word frequency effect. This is consistent with a number of models of language processing in which the word and the morpheme provide two independent levels of lexical representation, the word and the morpheme (e.g. Burani & Caramazza, 1987). Existing models, however, do not address the relative influence of the two levels of lexical representations on word recognition. Our results add to a growing body of evidence (e.g. Bertram, Schreuder, et al., 2000; Marslen-Wilson et al., 1994) suggesting that the relative importance of whole word and morphemic processes in recognition is a function of the linguistic properties of individual lexical items. It is noteworthy, however, that in the current study the effects of base morpheme frequency were markedly smaller than those of word frequency. This suggests that although both whole-word and morphemic processes are important determinants of response times, word representations dominate lexical processing.

The current study provides convincing demonstrations that affix productivity has a robust and important influence on the recognition of derived words. However, further research is needed before productivity can take its place with semantic transparency (Marslen-Wilson et al., 1994) as an important determinant of lexical representation and processing in languages like English. First, although we have demonstrated the importance of affix productivity for lexical representation, productivity was treated as a categorical variable rather than a linear measure. This means that the design was not sensitive to the possibility of graded differences of representation between words with more or less productive suffixes. Further research is needed to determine whether productivity is a categorical or a graded phenomenon. This would require regression analyses of responses to derived words of varying productivity and would require the development of tractable, ecologically-valid continuous measures of affix productivity from corpora or behavioural measures in healthy, linguistically naïve volunteers. Second, the question of whether productivity can be separated from affix homonymy may require further investigation. Although homonymy was not the main focus of this research, post hoc analyses suggest that the base morpheme frequency effects in Experiment 1 were not influenced by affix homonymy, contrary to the claims of Bertram and colleagues (2000). However, in English derivationally non-productive suffixes are also more likely to be homonymic, making it difficult to find materials to test the effects of productivity and affix homonymy separately. In addition, the relative frequencies of each homonymic use may be important (Bertram, Schreuder, et al., 2000) and the degree of pseudo-affixation may also play a role (Laudanna, Burani, & Cermele, 1994). Separating the relative influence of affix homonymy and productivity will be challenging since these two properties of affixes are often correlated. One important method for gaining leverage on this question will be through cross-linguistic experiments, in which the influence of these two variables can be more carefully separated.

One potential confound in the current study is that the complex non-word fillers were all non-word bases with real suffixes, e.g. *brishly*. Taft (2004) found a facilitatory effect of base morpheme frequency on response times to inflected words when the fillers were constructed in a similar manner to those in the current study, with non-word bases and existing suffixes. An inhibitory effect was found when the complex non-words were novel combinations of existing bases and suffixes. This raises the possibility that in the current study, effects of base morpheme frequency might reflect the type of complex non-word fillers used and word non-word classification might have been achieved by focussing on the bases alone. However, there were robust effects of word-form frequency in all three experiments of the current study and these were always substantially larger than the effects of base frequency. If participants had focussed on the bases of the derived words, effects of base morpheme frequency should have been at least as large as those of word-form frequency (cf. Taft, 2004, Experiment 1). Moreover, if such a strategy had been used, it is unclear to us why this would be the case for words with more productive suffixes but not for those with less productive suffixes. Thus although non-word context can influence morphemic frequency effects, we do not feel that the results of the experiments in this study were unduly influenced by our choice of non-word foils.

In conclusion, frequency effects in lexical decision continue to provide important constraints on theoretical models of the language system. The current findings of facilitatory effects of base morpheme frequency for productively suffixed derived words in English suggest that language representation is sensitive to the regularities provided by productive derivational morphology. Future studies of the role of morphemic variables on lexical representation and processing should strive to assess the influence of this important lexical variable.

Appendix A. Stimuli, Experiment 1

abruptly	campaigner	donation	gambler
absurdity	cancellation	doubtful	gardener
acceptance	captivity	drainage	global
accountant	chaotic	drinker	glorious
adaptable	chastity	effortless	grimly
adjustment	coastal	endurance	harassment
agility	coldness	energetic	heater
allegation	comical	enjoyment	helper
amusement	commuter	entertainment	herbal
assassinate	conditional	eruption	hermitage
assertion	confession	evasive	historian
athletic	continuation	excessive	hopeful
attacker	contradiction	expectation	hospitalise
banishment	controller	exploitation	hostility
beggar	costly	feeder	hugely
believer	councillor	fertilise	hunter
betrayal	cowardice	fighter	ignition
bleakly	decoration	flattery	imitation
bluntly	densely	flirtation	involvement
bridal	deportation	follower	irrigation
brutal	destroyer	forgiveness	killer
builder	detachment	friendship	kitchenette
burglary	dietary	frustration	learner
buyer	dismissal	futility	lender
listener	orphanage	saver	strangeness
locally	parental	scavenger	structural
loser	pilgrimage	scenic	supportive
magnetic	planner	sculpture	supremacy
membership	politely	seduction	swimmer
merciful	prescription	seizure	systematic
merrily	prestigious	serenity	taxable
mileage	priceless	seriousness	temptation
ministerial	privately	shrinkage	theorise
modernise	pronunciation	similarity	timidity
moisture	prosecution	simplify	trader
momentary	provincial	sincerity	treatment
monthly	publicly	smoker	troublesome
musician	readiness	snobbish	vanity
narration	resignation	socially	vicarage
nobility	retrieval	speaker	weaponry
noticeable	revision	specially	weaver
obscenity	revival	standardise	weekly
obscurity	richness	starvation	winner
occurrence	rotation	steepness	wreckage
orderly	sanity	sterilise	wrongly

Appendix B. Experiment 1, productivity classification (affixes are classified as “more or “less productive according to each of the three classification procedures)

Affix	Prod 1	Prod 2	Prod 3
able	More	More	More
ation	More	More	More
er	More	More	More
ful	More	More	More
ish	More	More	More
less	More	More	More
ly	More	More	More
ness	More	More	More
ify	Less	More	More
ment	Less	More	More
ship	Less	Less	More
ise	Less	Less	More
age	Less	Less	Less
al	Less	Less	Less
ance	Less	Less	Less
ant	Less	Less	Less
ary	Less	Less	Less
ate	Less	Less	Less
cy	Less	Less	Less
ence	Less	Less	Less
ery	Less	Less	Less
ette	Less	Less	Less
ian	Less	Less	Less
ic	Less	Less	Less
ice	Less	Less	Less
ion	Less	Less	Less
ious	Less	Less	Less
ity	Less	Less	Less
ry	Less	Less	Less
some	Less	Less	Less
ure	Less	Less	Less

Appendix C. Stimuli, Experiment 2

More productive affixes			
abruptly	entertainment	planner	amazement
adaptable	expectation	politely	announcement
allegation	exploitation	privately	assessment
amusement	feeder	pronunciation	attachment
attacker	fighter	readiness	awareness
banishment	flirtation	resignation	baker
bleakly	follower	richness	baldness
bluntly	forgiveness	saver	blindness
buyer	gambler	scavenger	briskly
campaigner	gardener	seriousness	composer
cancellation	grimly	smoker	dampish
commuter	harassment	snobbish	dancer
densely	heater	specially	deafness
controller	hopeful	starvation	dissenter
costly	hugely	steepness	excitement
councillor	hunter	strangeness	exploration
decoration	involvement	swimmer	faithful
deportation	killer	taxable	fiercely
destroyer	lender	trader	harmless
detachment	listener	treatment	hollowness
doubtful	locally	weaver	daftness
effortless	merrily	wrongly	legally

Appendix C (*continued*)

enjoyment	noticeable	alertness	literally
madness	offender	shameful	slimness
neatness	predictable	sickness	publicly
novelist	mildness	singer	weirdness
Less productive affixes			
drainage	acceptance	confession	agility
orphanage	endurance	contradiction	captivity
pilgrimage	accountant	donation	chastity
shrinkage	beggar	eruption	futility
vicarage	burglary	frustration	hostility
wreckage	dietary	ignition	nobility
betrayal	momentary	imitation	obscenity
bridal	assassinate	irrigation	obscurity
brutal	supremacy	prescription	sanity
coastal	occurrence	prosecution	serenity
comical	flattery	revision	similarity
conditional	kitchenette	rotation	sincerity
dismissal	historian	seduction	timidity
global	musician	glorious	vanity
herbal	athletic	prestigious	excessive
ministerial	chaotic	fertilise	supportive
parental	energetic	hospitalise	adjustment
provincial	magnetic	modernise	weaponry
retrieval	cowardice	sterilise	friendship
revival	simplify	theorise	membership
structural	assertion	absurdity	troublesome
moisture	certainty	forestry	rejection
sculpture	continuous	harden	rental
seizure	darken	recovery	storage
attraction	delivery	reflection	straighten
advisory	denial	refusal	warmth
burial	exposure	regional	widen

Appendix D. Stimuli, Experiment 3

accountant	darken	blockage	usage
conditional	forestry	complainant	sweeten
herbal	beggar	deepen	toughen
weaponry	coastal	fishery	avoidance
certainty	bridal	assessor	freshen
supportive	acceptance	inhibitory	cookery
membership	occurrence	shorten	corrective
supremacy	seizure	scenic	exhibitor
energetic	dietary	fragmentary	whiten
burial	kitchenette	kingship	progression
continuous	musician	parentage	dosage
troublesome	sadden	roughen	greenery
momentary	worrisome	riotous	creator
drainage	detector	township	brighten
shrinkage	lectureship	lighten	distributor
retrieval	olden	assertive	mileage
obscurity	sicken	creditor	quicken
denial	reflector	elector	thicken
similarity	connective	disciplinary	nunnery
structural	statuette	angelic	trickery
refusal	admittance	joyous	contributor
harden	blacken	prohibitory	inventor
widen	rockery	treasonous	postage
friendship	deletion	debtor	observant
informant	dilution	breakage	absorbent
authorship	quieten	baggage	completion

(*continued on next page*)

Appendix D (*continued*)

citizenship	thuggery	protector	modernity
waken	metallic	murderous	oddity
prevention	royalty	resistant	secretive
alphabetic	historic	claimant	

References

Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language, 40*, 41–61.

Baayen, H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37*, 94–117.

Baayen, R. H., Feldman, L. B., & Schreuder, R. (2006). Morphological influences on the recognition of monosyllabic monomorphemic words. *Journal of Memory and Language, 55*, 293–313.

Baayen, R. H., & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics, 29*(4), 801–844.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database (CDROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.

Bauer, L. (1983). *English word formation*. Cambridge, UK: Cambridge University Press.

Bertram, R., Baayen, R. H., & Schreuder, R. (2000). Effects of family size for complex words. *Journal of Memory and Language, 42*, 390–405.

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: Learning, Memory and Cognition, 26*, 489–511.

Bradley, D. (1979). Lexical representation of derivational relation. In M. L. Kean (Ed.), *Juncture*. Cambridge, MA: MIT Press.

Broadbent, D. E. (1967). Word frequency effect and response bias. *Psychological Review, 74*, 1–15.

Burani, C., & Caramazza, A. (1987). Representation and processing of derived words. *Language and Cognitive Processes, 2*, 217–227.

Burani, C., Salmaso, D., & Caramazza, A. (1984). Morphological structure and lexical access. *Visible Language, 18*(4), 342–352.

Clahsen, H., Eisenbeiss, S., & Sonnenstuhl-Henning, I. (1997). Morphological structure and the processing of inflected words. *Theoretical Linguistics, 23*(3), 201–250.

Colé, P., Beauvillain, C., & Segui, J. (1989). On the representation and processing of prefixed and suffixed derived words: A differential frequency effect. *Journal of Memory and Language, 28*, 1–13.

Colombo, L., & Burani, C. (2002). The influence of age of acquisition, root frequency, and context availability in processing nouns and verbs. *Brain and Language, 81*, 398–411.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, NJ: Erlbaum.

De Jong, N. H., Schreuder, R., & Baayen, R. H. (2000). The morphological family size effect and morphology. *Language and Cognitive Processes, 15*, 329–365.

Forster, K. I. (1976). Accessing the mental lexicon. In R. J. Wales & E. Walker (Eds.), *New approaches to language mechanisms* (pp. 257–287). Amsterdam: North-Holland.

Forster, K. I., & Forster, J. C. (2003). A Windows display program with millisecond accuracy. *Behavior Research Methods Instruments and Computers, 35*, 116–124.

Giraudo, H., & Grainger, J. (2000). Effects of prime word frequency and cumulative root frequency in masked morphological priming. *Language and Cognitive Processes, 15*, 421–444.

Hauk, O., Davis, M. H., Ford, M., Pulvermüller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *Neuroimage, 30*(4), 1383–1400.

Howes, D. H., & Solomon, R. L. (1951). Visual duration threshold as a function of word-probability. *Journal of Experimental Psychology, 41*, 401–410.

Kempley, S., & Morton, J. (1982). The effects of priming with regularly and irregularly related words in auditory word recognition. *British Journal of Psychology, 73*, 441–454.

Laudanna, A., Burani, C., & Cermele, A. (1994). Prefixes as processing units. *Language and Cognitive Processes, 9*, 295–316.

Lorch, R. F., Jr., & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*, 149–157.

Marchand, H. (1969). *The categories and types of present-day English word-formation* (2nd ed.). Munich: C.H. Beck.

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.

Matthews, J. N. S., Altman, D. G., Campbell, M. J., & Royston, P. (1990). Analysis of serial measurements in medical research. *British Medical Journal, 300*, 230–235.

McCormick, S. F., Rastle, K., & Davis, M. H. (2007). Is there a fete in fetish? Effects of orthographic opacity on morpho-orthographic segmentation in visual word recognition. *Journal of Memory and Language*. doi:10.1016/j.jml.20.

Meunier, F., & Segui, J. (1999). Frequency effects in auditory word recognition. *Journal of Memory and Language, 41*, 327–344.

Morton, J. (1969). Interaction of information in word recognition. *Psychological Review, 76*, 165–178.

Niswander, E., Pollatsek, A., & Rayner, K. (2000). The processing of derived and inflected suffixed words during reading. *Language and Cognitive Processes, 15*, 389–420.

Oxford English Dictionary on CD-ROM (1992). Oxford University Press, Oxford.

Plag, I. (2004). Productivity. In B. Aarts & A. McMahon (Eds.), *Handbook of English linguistics* (pp. 537–576). Oxford: Blackwell.

Schreuder, R., & Baayen, R. H. (1995). Modeling morphological processing. In A. Friederici (Ed.), *Language comprehension: A biological perspective* (2nd ed., pp. 101–132). Berlin: Springer.

Schreuder, R., & Baayen, R. H. (1997). How complex simplex words can be. *Journal of Memory and Language, 41*, 327–344.

Sereno, J. A., & Jongman, A. (1997). Processing of English inflectional morphology. *Memory & Cognition, 25*, 425–437.

Stevens, J. P. (1999). *Applied multivariate statistics for the social sciences* (3rd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Taft, M. (1979). Recognition of affixed words and the word-frequency effect. *Memory & Cognition, 7*, 263–272.

Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes, 9*, 271–294.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology Section A – Human Experimental Psychology, 57*, 745–765.