Mechanisms for remembering roots versus affixes in complex words

Anne Pycha

*Corresponding author's address: Linguistics, University of Wisconsin, Milwaukee, Milwaukee, Wisconsin 53211, pycha@uwm.edu

Previous research has demonstrated that listeners remember low-frequency words (fob) through explicit recollection, but high-frequency words (money) through implicit familiarity (Joordens & Hockley 2000). We hypothesize that a similar asymmetry in remembering occurs in morphologically complex words (bleakish), where root frequency (bleak) is always low relative to affix frequency (ish). In our experiment, which modifies a technique developed by Goldinger, Kleider & Shelley (1999), participants hear both complex and simple words at study. At test, they hear old and new words in which a portion of the stimulus is overlaid with background noise. For complex words, the masked portion is either the root or the affix; for simple words (polish), it is the corresponding pseudo-morpheme. Participants rate the subjective loudness of the noise on a five-point scale. Results do not indicate any difference in ratings between old and new words, although participants did give significantly softer ratings to root noise for prefixed words, versus suffixed words. Thus, while listeners exhibit no apparent asymmetry in remembering roots and affixes, they do exhibit an asymmetry in perceiving them, suggesting a potential explanation for the general resistance of prefixed roots to phonological alternations.
INTRODUCTION

Memories and Speech Processing

The act of remembering plays an important role in speech processing, because listeners must compare the incoming speech stream with their memories for words that they have previously heard. In the model formalized by Halle and Stevens (1962) and elaborated upon by Poeppel & Monahan (2010), the listener uses basic perceptual processes to generate initial guesses about what she might be hearing (sounds like “gua”). These “guesses” are actually stored representations in memory ([GUAVA]?). The putative acoustic realizations of the stored representations feed back, in a loop, where the listener uses them to process the continuing stream of perceptual input (prediction: I’ll hear “va” next). A key idea, then, is that stored representations – that is, memories – have predictive value during word recognition, a notion also proposed in the TRACE model of word recognition (McClelland & Elman 1986).

Memories, however, are subjective experiences, and there is more than one way to remember. We may accurately remember an event as a set of features that we previously perceived, or we may falsely remember it as a combination of features that we have both perceived and imagined (Johnson et al. 2012). We may explicitly recollect an event in all of its detail, or we may implicitly acknowledge its familiarity (Yonelinas & Jacoby 2012). The distinctiveness of an event plays a key role here. Distinct events – for example, a misspelled word like dreem (Cortese et al. 2008) – are consistently less likely to produce false memories, and more likely to be remembered with explicit recollections, than indistinct events. Furthermore, our memories for distinct events contain more contextual information – for example, I heard “dog”, and I also saw a picture of a dog – than do our memories for indistinct events (Israel & Schacter 1997). Finally, people remember distinct events better overall: in their review of 80 previously published reports, Glanzer & Adams (1985) found that, in recognition tasks, distinct events consistently produced higher hit rates and lower false alarm rates than non-distinct events. This finding held for unique vs. typical faces, pictures vs. words, concrete vs. abstract words, real words vs. nonsense words, and infrequent vs. frequent words.

Roots and Affixes

In the current experiment, we pursue the hypothesis that people remember affixes and roots differently. In complex words such as reshred and clownship, for example, we investigate whether listeners form different types of memories for the prefix re- and the suffix -ship versus the roots shred and clown. We had two motivations for this hypothesis. The first motivation is somewhat specific, and comes from previous work on word memories. Many studies have already established that people remember high-frequency and low-frequency words differently (e.g., money vs. job), which researchers attribute to different processes of remembering (Joordens & Hockley 2000). According to these accounts, we have many pre-existing experiences with high-frequency words, so they constitute less distinctive events that we remember through a process of implicit familiarity. By contrast, we have only a few previous experiences with low-frequency words, so they constitute highly distinctive events that we remember through a process of explicit recollection.

Despite the fact that they cannot occur in isolation, individual affixes such as re- and -ship have relatively high token frequencies; they come from a small, closed class of morphemes that we draw upon over and over again in order form new words. Individual roots such as shred and clown, by comparison, have relatively low token frequencies; they come from a large, open class of morphemes from which we can select the specific meaning we need. Given these frequency differences between roots and affixes, the same logic that applies to the memories of whole words could also apply to individual morphemes. That is, we hypothesize that affixes constitute less distinct events than roots, and that this difference impacts the memories we form of them.

The second motivation for the current experiment is broader, and comes from an asymmetry in cross-linguistic word patterns. In general, morpheme boundaries are very active sites for phonological alternations. Root segments, for example, frequently trigger changes in prefixes (im-possible, in-destructible) or in suffixes (thought-s, wish-es). Similarly, suffix segments frequently trigger changes in roots (Italian dok-t-e → do-te ‘led’, Saltarelli 1970). But there is a gap in this paradigm: prefix segments only rarely trigger changes in roots (one exception comes from Luganda, m-báñ-a → m-máñ-a ‘I demand payment’, Hyman & Katamba 1999: 397). A diverse set of processes observe this alternation asymmetry, including local consonant assimilation, vowel harmony, consonant harmony, and vowel elision (Hyman 2008). Many previous accounts of this asymmetry have emphasized the primacy of root...
morphemes during speech processing (e.g., Hawkins & Cutler 1988), but this does not explain why languages tolerate root changes that originate from suffixes only, and not from prefixes. If we are going to wreak havoc on the root’s role in word recognition, why does it matter where the havoc comes from? We hypothesize this asymmetry may be linked to different memories that people form not just of roots versus of affixes, but of more specifically, of prefixes versus suffixes.

Experimental Approach

Our research question presented a methodological problem. Unlike words, affixes cannot be presented in isolation, making it difficult to assess whether listeners treat them differently than roots. To solve this problem, we adapted a perceptual illusion technique. Previous studies have shown that if people create a memory for a stimulus, they perceive that stimulus differently later on. In Jacoby et al. (1989), for example, participants gave higher fame estimates to non-famous names if they had seen those names earlier in the experiment. In Jacoby et al. (1988), participants gave lower volume ratings to background noise if they were listening to sentences they had heard earlier.

Goldinger, Kleider, & Shelley (1999) replicated this finding using individual words. For example, at study, participants heard river, symbol, bicep and other words. At test, they heard a mix of old and new words presented with background noise. Participants indicated that old words, like river, had softer background noise, as long as they were presented in the same voice. This is a perceptual illusion, because noise volume remained constant across old and new. The use of background noise offered a solution to our problem, because we could overlay noise on just one morpheme of a complex word while presenting the other morpheme in the clear, allowing us to assess the effect of noise on individual roots and affixes while still presenting complex words in their entirety. We therefore adapted the techniques of Goldinger, Kleider & Shelley (1999)’s Experiment 1.

METHOD

Overview

As schematized in Table 1, participants completed a study phase in which they listened to seventeen affixed words presented in the clear, such as reshred and clownship. They then completed a test phase in which they listened to thirty-four words partially overlaid with white background noise, and rated the loudness of the noise on a scale from 1 (softest) to 5 (loudest). At test, approximately half of the words were old, having been already presented at study, and half were new. Furthermore, approximately half of the words had noise overlaid only the root, while half had noise only on the affix. The actual noise occurred at one of three different signal-to-noise ratios, but participants offered a subjective rating from 1 (softest) to 5 (loudest). The key questions were a) whether listeners would give different noise ratings to old versus new words, and, b) if this perceptual illusion were present, whether listeners were more prone to experience it on roots versus affixes, or on prefixes versus suffixes.

<table>
<thead>
<tr>
<th>Study: Listen and identify</th>
<th>Test: Rate loudness of noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>reshred</td>
<td>babedom</td>
</tr>
<tr>
<td>clownship</td>
<td>bleakish</td>
</tr>
<tr>
<td>bewhirl</td>
<td>unehalk</td>
</tr>
<tr>
<td>bleakish</td>
<td>bewhirl</td>
</tr>
<tr>
<td>...</td>
<td>fudgeless</td>
</tr>
<tr>
<td></td>
<td>reshred</td>
</tr>
<tr>
<td></td>
<td>clownship</td>
</tr>
</tbody>
</table>

As a control, participants also completed identical study and test phases with monomorphemic words such as regret, which contains the pseudo-prefix re-, and worship, which contains the pseudo-suffix -ship. So as not to encourage participants to treat these pseudo-affixes as real affixes, we presented these control words in a separate session prior to the target words.
Word lists

The target stimuli were created from a list of morphologically complex words that we coined, such as *reshred* and *clownship*, while aiming for three key characteristics: a) low root frequency, b) high parsability into component morphemes, and c) matched parsability across prefixed and suffixed conditions.

Although familiar, the roots *shred* and *clown* have relatively low frequencies (three occurrences per million words, according to the CELEX database, corresponding to a log frequency of 1.48), maximizing their distinction with affixes such as *re-* and -*ship*, which have relatively high frequencies. According to the logic outlined above, this disparity in frequency counts, which we imposed on all stimuli by using roots with a log frequency of 2.00 or less, should increase the likelihood that listeners treat roots and affixes differently, a desirable scenario for this experiment.

The words *reshred* and *clownship* are also highly parsable into their component morphemes, again facilitating a comparison between those morphemes. We imposed high parsability in all stimuli by manipulating relative frequency and phonotactic transitions. As Hay (2006) argued, for derived forms which are more frequent than their roots (*illegible*, for example, is more frequent than *legible*), listeners are more likely to use whole-word access, while for derived forms which are less frequent than their roots (*illiberal* is less frequent than *liberal*), listeners are more likely to parse the word into individual morphemes. In our stimuli, the derived forms are newly-coined and therefore have a frequency of zero, so they are always less frequent than their roots. Low probability of phonotactic transitions across morphemes boundaries also facilitates morpheme parsing (Hay 2006); we constructed stimulus words so that this value was as close to zero as possible (0.0002 for [i-f] transition in *re-shred*, 0 for [n-f] in *clownship*; Vitevitch & Luce 2004).

Finally, the prefixed word *reshred* and the suffixed word *clownship* are also matched for parsability, helping to ensure that any difference in how listeners treat these words arises from their prefixed or suffixed status per se, and not from extraneous differences. We imposed matching across prefix and suffix conditions by using Hay & Baayen’s (2002) “type parsing ratio”, which indicates, for a given affix, the proportion of derived words which listeners are likely to parse into component morphemes. The prefix *re-* for example, has a ratio of 0.68, which means that among the words derived from this (such as *redo, rework*, *revisit*, *rehabilitate*, etc.), listeners are likely to parse 68% of them into morphemes, while processing the remaining 32% as whole words. The matched suffix -*ship* has a ratio of 0.62, suggesting a similar probability of listener parsing. Each prefixed word in the stimulus list was matched as closely as possible with a suffixed word, creating balance across conditions; the mean type parsing ratio was 0.65 for prefixes (ranging from 0.43 to 0.93) and 0.65 for suffixes (ranging from 0.42 to 0.92).

Hay & Baayen (2002) provide type parsing ratios for 26 prefixes. In order to minimize variation in the raw duration of target morphemes, we restricted our stimuli to monosyllabic affixes, which left us with 19 prefixes to work with. Of these, *em-* was eliminated because it is a phonological variant of *en-* and *cross-* was eliminated because no suffix with a comparable type parsing ratio was available. The resulting seventeen prefixes were matched, based on the criteria above, to seventeen suffixes, which we concatenated to monosyllabic, low-frequency roots in order to coin thirty-four complex, disyllabic words. No root or affix was included on the list more than once.

We also selected thirty-four control words which were mono-morphemic, disyllabic, and contained pseudo-affixes to match each of the real affixes in the target words (*re-gret, wor-ship*, etc). Only a small handful of words in English meet these criteria, so we were unable to control their lexical characteristics any further.

Stimulus preparation

A linguistically-trained speaker of American English, who was not aware of the purpose of the experiment, recorded each word in a sound-proof booth. With the Praat program (Boersma & Weenink 2012), we used waveforms and spectrograms to segment each recorded word into root and affix portions. With the Akustyk program (Plichta 2012), we added white noise to either the root or the affix at one of three signal-to-noise ratios: soft (+24 dB), medium (+17 dB), or loud (+10 dB). Thus, for each word such as *reshred*, six stimuli were created: three with noise on the root (*reshred* at soft, medium, and loud S/N ratios) and three with noise on the affix (*reshred* at soft, medium, and loud S/N ratios).

Design

During the study phase, listeners heard a list of 17 words spoken over headphones, which were randomly selected and ordered for each participant from the pool of 34 complex words. For each word, they saw two buttons on the screen (e.g., *clownship* and a foil such as *basis*), and they identified the word they heard by clicking on the
appropriate button. During the test phase, listeners heard all 34 words spoken over headphones by the same speaker. Each word was partly overlaid with background noise, and participants rated the loudness of the noise by clicking on a button from 1 (softest) to 5 (loudest). The order of words, placement of noise on a word’s root versus its affix, and S/N ratio (soft, medium, loud) of the noise, were all randomized for each participant. This study-test paradigm was preceded by an identical study-test paradigm for the thirty-four simple control words.

Participants

One hundred twenty-seven adult members of the University of Wisconsin, Milwaukee community participated in the experiment, which lasted approximately 15 minutes, and received either course credit or $5 as compensation. Seventeen subjects were excluded from the analysis because they were not native speakers of English, leaving 110 subjects who were asked to make 68 judgments each (34 complex words, 34 simple control words). Of the 7480 judgments collected, 75 were blank cells corresponding to stray mouse clicks, so our final analysis included 7405 judgments.

RESULTS

Results did not confirm the hypothesis. Participants rated the loudness of background noise equivalently for old and new words, and this null finding held for all study conditions, suggesting that stored memories did not alter perceptual fluency. However, the results also show that participants rate noise on prefixed versus suffixed words very differently, suggesting that prefixed words (and in particular, roots preceded by a prefix) provide more perceptual clarity overall than suffixed words do.

We used stepwise regression to identify the best-fitting model for the data (with the `step()` function from the stats package for R), entering the following predictors and interactions: Word status x Noise location x Affix type x Morpheme type x S/N ratio. We then used proportional odds logistic regression to calculate the regression coefficients and statistics (with the `polr()` function from the MASS package for R). Results of the model are summarized in Table 2.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noise location (root vs. affix)</td>
<td>-0.69</td>
<td>0.062</td>
<td>-1.11</td>
<td>0.26</td>
</tr>
<tr>
<td>Affix type (prefix vs. suffix)</td>
<td>0.36</td>
<td>0.099</td>
<td>3.61</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Morpheme type (real vs. pseudo)</td>
<td>-1.02</td>
<td>0.063</td>
<td>-16.20</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>S/N ratio (soft vs. medium)</td>
<td>1.91</td>
<td>0.084</td>
<td>22.87</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>S/N ratio (medium vs. loud)</td>
<td>4.06</td>
<td>0.087</td>
<td>46.44</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Noise location &amp; Affix type</td>
<td>-0.59</td>
<td>0.088</td>
<td>-6.72</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Affix type &amp; Morpheme type</td>
<td>0.30</td>
<td>0.088</td>
<td>3.45</td>
<td>&lt; 0.01</td>
</tr>
<tr>
<td>Affix type &amp; S/N ratio (medium)</td>
<td>-0.01</td>
<td>0.108</td>
<td>-0.10</td>
<td>0.92</td>
</tr>
<tr>
<td>Affix type &amp; S/N ratio (loud)</td>
<td>-0.28</td>
<td>0.108</td>
<td>-2.60</td>
<td>&lt; 0.01</td>
</tr>
</tbody>
</table>

Main effects

Word status (old versus new) was not included in the best-fitting model, so there is no indication that participants rated noise differently if they had previously heard the word at study, contrary to our initial hypothesis. Noise location was included in the model; its main effect was not significant, but it interacted with affix type.

Affix type, morpheme type, and S/N ratio were all significant predictors. As indicated by the respective coefficients, listeners were more likely to give a) louder noise ratings to suffixed words than to prefixed words, b) softer noise ratings to pseudo-morphemic words than to multi-morphemic words, and c) louder ratings when the S/N ratio was medium (compared to soft) or loud (compared to medium).
Interactions

The interaction between noise location and affix type was significant, as illustrated in Figure 1, which shows the cumulative probabilities of noise ratings at Levels 1 through 5. For prefixed words, participants were equally likely to assign a given rating when they heard noise on the root (\textit{reshred}, \textit{regret}) or on the affix (\textit{reshred}, \textit{regret}). But the same pattern did not hold for suffixed words, where participants were more likely to assign a louder rating when they heard noise on the root (\textit{clownship}, \textit{worship}), but a softer rating when they heard noise on the affix (\textit{clownship}, \textit{worship}).

The interaction between affix type and morpheme type was also significant. In multi-morphemic words, participants were equally likely to assign a given noise rating to prefixed versus suffixed words, but in pseudo-morphemic words, participants were more likely to assign soft noise ratings to prefixed words (\textit{regret}) and louder ratings to suffixed words (\textit{worship}). This effect held regardless of whether the noise occurred on the root or the affix.

Finally, the interaction between affix type and S/N ratio was significant, as illustrated in Figure 2. When the noise occurred at the loud S/N ratio, participants were equally likely to assign a given noise rating to prefixed versus suffixed words. But when the noise occurred at the soft or medium S/N ratios, participants were more likely to assign softer noise ratings to prefixed words (\textit{reshred}, \textit{regret}), and higher noise ratings to suffixed words (\textit{clownship}, \textit{worship}), again regardless of where the noise occurred.
FIGURE 2. Interaction between S/N ratio and affix type. The x-axis indicates the S/N ratio in the stimulus: soft (+24 dB), medium (+17 dB), or loud (+10 dB). The y-axis indicates the cumulative probability of listeners assigning a particular subjective loudness rating to the noise, from Level 1 (softest) to Level 5 (loudest). Comparison of the data for prefixed words (left panel) and suffixed words (right panel) shows that listeners responded differently to prefixed versus suffixed words in the soft and medium conditions, compared to the loud condition.

DISCUSSION

This study found no evidence of a perceptual illusion effect; that is, we have no evidence that the recent memory of a word enhances its perceptual clarity. Thus, we cannot confirm our initial hypothesis that listeners remember roots and affixes differently. Our results differ from those of Goldinger, Kleider & Shelley (1999), who did find a perceptual illusion effect. They also differ from results reported for high- versus low-frequency words, which listeners remember through different mechanisms (Joordens & Hockley 2000).

One key difference between our current experiment and previous work is that our investigation focused primarily on multi-morphemic words, rather than simple words. It is possible that the mechanisms for remembering and perceiving such words differ significantly from those for simple words. If this difference played a crucial role, however, we would still expect to see a perceptual illusion effect for the control words included in the experiment (regret, worship), but this was not evident in our results.

Another key difference is that our investigation focused on parts of words, rather than entire words. Specifically, whereas Goldinger, Kleider & Shelley (1999) presented test words covered entirely with background noise, we presented test words covered only partially, on the root or affix. Because the remainder of the word occurred in the clear, participants perhaps did not rely very heavily on stored memories during word recognition, instead relying on basic perceptual processes that operate in the same way regardless of whether a word is old or new.

Although this experiment did not uncover differences in memory, it did reveal interesting differences in the basic perception of roots versus affixes, and these differences could have potentially intriguing consequences for our understanding of the phonological alternation asymmetry (recall that roots and suffixes trigger changes, but prefixes generally do not). Most notably, our results show that for prefixed words, listeners respond to noise similarly regardless of whether it occurs on the prefix or the root; for suffixed words, however, listeners perceive noise
occurring on the root as significantly louder than noise occurring on the affix. Reading off the data from Figure 1, we can arrange word types into the following hierarchy:

\[
\text{Increasing probability of noisier ratings } \rightarrow \\
\text{clownship} \ll \text{reshred} \ll \text{clownship}
\]

On the assumption that softer ratings correlate with enhanced perceptual clarity, one interpretation of our results is that the perceptual clarity of roots depends crucially upon word structure. Listeners perceive roots pretty robustly when they are preceded by a prefix, but otherwise, listeners perceive roots more tenuously. Such a finding runs counter to the oft-cited idea that listeners perceive words optimally when the root occurs first (see Hyman 2008 and references cited therein). If this finding were to hold up in further experiments, and in other languages, it could suggest an intriguing solution to the problem of the alternation asymmetry: post-prefix roots are perceptually clear, therefore uniquely resistant to listener mis-perception and the development of alternations (Ohala 1990), while pre-suffix roots enjoy no such special status. Of course, such an account would still have to explain why post-root suffixes regularly undergo alternations, despite their presence at the top of this hierarchy.

Our results also suggest an overall perceptual advantage for prefixed words regardless of noise location, although it is difficult to pin down the implications of this finding. Participants were overall more likely to assign softer ratings to prefixed words, but, as the interactions in our statistical model revealed, this effect occurs in the soft and medium S/N ratios, and was largely absent from the loud S/N ratio, which may simply indicate a heightened sensitivity to difference when actual noise levels are softer. Curiously, this perceptual advantage appears to originate largely from simple pseudo-morphemic words, and is mostly absent from complex multi-morphemic words, a finding for which we have no straightforward explanation, except to suggest that pseudo-prefixes may be more likely to undergo obligatory decomposition than pseudo-suffixes (e.g., Rastle, Davis, & New 2004).

In sum, future investigations into the hypothesized memory differences between roots and affixes should probably tax listeners’ memories more heavily than we did in the current experiment. But as an investigation into perceptual differences between roots and affixes, we report some interesting, although admittedly post-hoc, results: listeners do not perceive all roots alike. Rather, the perceptual robustness of a root seems closely linked to the morphological structure in which it finds itself, a finding which, if upheld, may eventually help explain a puzzling asymmetry in phonological typology.

ACKNOWLEDGMENTS

I thank Dave Armstrong, Ivan Ascher, Garry Davis, Christian DiCanio, Fred Eckman, Nicholas Fleisher, Jami Eaton, Jenny Lederer, John Kingston, Lydia Murphy, Hanyong Park, Jae Young Song, Suzanne van der Feest, Kari Whittenberger-Keith, and Carolyn Zafra. I also thank the creators of the resources that I relied upon, at no cost to me: Akustyk, CELEX, Neighborhood Activation Database, Phonotactic Probability Calculator, Praat, and R. This project was funded by the University of Wisconsin, Milwaukee.

REFERENCES


