LEXICAL ACTIVATION BY ASSIMILATED AND REDUCED TOKENS

M. Louise Kelly, Ellen Gurman Bard*†, Catherine Sotillo†*

Department of Linguistics*, Human Communication Research Centre†

University of Edinburgh

ABSTRACT

Running speech contains abundant assimilated and phonologically reduced tokens, but there is considerable debate about how such varied pronunciations disrupt access to the corresponding words in listeners' mental lexicons. While previous studies have examined the effects of carefully produced or electronically edited reductions, we present two experiments which compare cross-modal repetition priming for lexical decision by more reduced spontaneous forms and less reduced read forms of the same words uttered by the same speakers in the same phrases. Though less priming is found for the more reduced spontaneous tokens, both versions of words produce significant priming effects, whether the majority of stimuli are taken from spontaneous speech (Experiment 1) or from read speech (Experiment 2). Priming is more robust if the tokens themselves contain the context licensing the reduction.

1. INTRODUCTION

Recent work on word recognition from running speech has debated how much signal variation the listener can deal with and still successfully map speech onto a representation in the mental lexicon.

Some experimental evidence from priming and lexical decision paradigms support the notion that any phonological variation from the canonical pronunciation is deleterious to that word's recognition. Marslen-Wilson & Zwitzerlood [1] suggested that any alteration to the word initial phoneme caused severe disruption to the lexical activation system. They found that while *honey* will prime BEE, *noney* will not. They claimed that phonological (albeit on one phoneme) variation which alters a token by several features is enough to dissipate activation considerably.

Other work, by Marslen-Wilson, Van-Halen & Moss [2], had suggested that alteration of just one feature of a phoneme was enough to disrupt lexical access, if that phoneme was word initial: *dask* primes JOB much more weakly than *task* primes JOB. Andruski et al. [3] and Connine et al. [4] also reported that lexical entries were activated in proportion to the featural mismatch between the spoken token and the stored lexical representation. These studies found that when words were presented in isolation, any distortion to them was detrimental to their being recognized.

Further complications to the pattern of priming were observed by Lahiri and Marlsen-Wilson [5]. They proposed that regular phonological variation does not disrupt lexical activation because the stored representations are underspecified in such a way that the reduced and assimilated tokens will still match the entry in the mental lexicon.

Within sentence contexts, lexical access is generally found to be more robust. When presented with a leading context, distorted tokens can activate their stored targets. Gaskell & Marslen-Wilson [6] found that regular phonological variation did not block lexical access assimilated for assimilated tokens which followed a felicitous sentence onset: thus [lim] primed LEAN after *'The house was full of fussy eaters. Sandra would only eat…'*

Subsequent contexts also affected the outcome: [wɪkɪb] primed WICKED before a context licensing the assimilation (e.g. *prank*) but not before a phonologically unviable context (e.g. *game*). The researchers suggested that phonologically predictable variation (such as assimilation) often demands processes of phonological inference. It is not clear, however, whether the presence of a licensing context permits the desired conclusion or whether an impossible context blocks it. It is clear that listeners were attending to more than the altered segments. If phonological inference cannot be completed until such contexts are encountered, then phonological effects dependent on word-internal structure (e.g. weak syllable reduction as in *gazelles*) should permit lexical activation more readily than those licensed by following words (e.g. nasal final assimilation as in *Crane Bay*).

The importance of the rest of what listeners hear brings up questions about the nature of the stimuli as a whole. In general, these have been carefully prepared so that they are flawless except for the critical change. Some studies used word tokens pronounced by phonologically sophisticated experimenters so as to include a single intentional distortion to otherwise careful speech. Others used acoustically manipulated materials that contain artificial sounds intermediate between two canonical targets. While the materials help to address the hypotheses, they do not reflect the characteristics of natural running speech where precise articulation is relatively rare. More important, they may create an expectation on the listeners' part that they are hearing clear speech where phonetic gestures are largely canonical. Even if presented out of sequence, more natural materials might be expected to maintain lexical access at a robust baseline whenever regular phonological processes are at work.

In this paper we present two cross-modal repetition priming experiments on the effects of phonological reductions sampled from spontaneous speech. Because priming for lexical decision is not found for phonologically similar but semantically unrelated words (*fort*/FORTITUDE: see [7]) priming is taken to

indicate activation of the appropriate entry in the mental lexicon. We test the following hypotheses:

- Quality: more reduced tokens will prime more weakly.
- Robust access: Nonetheless priming will be robust both for the more assimilated and/or reduced tokens typical of spontaneous speech and for the less assimilated tokens of typical of read speech.
- Ambient context: The quality of ambient speech will alter priming with reduced priming for assimilated tokens in generally clear speech.
- Licensing: Words whose internal structures license reductions or assimilations will prime more effectively than those whose reductions are dependent on external context.

2. EXPERIMENT 1

2.1. Method

2.1.1. Materials

All word tokens were landmark names excerpted from the stereo digital recordings of the HCRC Map Task Corpus [8]. The 60 Spontaneous tokens were single words taken from initial mentions of landmark names made by speakers of Southern Scottish English while negotiating the route-communication task. The 60 read tokens of these words were produced by the same speakers recording a list of multi-word landmark names (e.g. *Crane Bay, pebbled shore*). Both read and spontaneous tokens gave opportunities for assimilation.

Thirty-six read-spontaneous pairs of tokens had been produced in environments that offered the opportunity for assimilation or deletion: these were now excised. Twenty-four of these pairs were potential nasal place-of-articulation assimilations (e.g. ***pine** grove*); 12 were word final d-deletions (e.g. ***old** mill*). A further 24 contained their own conditioning environment: 12 were potential initial weak syllable reductions (***flamingo***) and 12 were potential non-initial weak syllable reductions (***buffalo***). Forty-eight further landmark words acted as fillers. Spontaneous tokens of the fillers were used in Experiment 1.

Each word was twinned with a length and frequency matched partner (e.g. ***train***, ***stone***). Excerpted single word tokens (e.g. ***train***) served as auditory primes. All words appeared as both auditory and visual stimuli and as primes and controls.

Visual stimuli were either

1. the word just heard, (e.g. auditory ***train*** - visual TRAIN)
2. a non-word formed from that word by altering one character, (e.g. ***train*** - TRAIX)
3. the twinned real word (***train*** - STONE)
4. or a non-word formed by altering one character of the twinned word (***train*** - STOCE).

Independent measures of each token's degree of reduction or assimilation [9] showed significant differences between read and spontaneous tokens. The latter were shorter in milliseconds and in standardized duration, less intelligible to naïve listeners asked to identify them, and judged to be more assimilated by phoneticians given forced choice tests. The difference between the more reduced spontaneous tokens and the less reduced read tokens formed our principal comparison.

2.1.2. Subjects and Procedure

Subjects were native speakers of English resident in Scotland with corrected to normal vision and no known hearing loss. They heard single auditory stimuli over headphones. A visual word appeared on screen automatically at the offset of the auditory item. Subjects were instructed to make lexical decision judgements to the visual stimuli. Subjects encountered each word only once and were exposed to equal numbers of primed and non-primed words and non-words. Seventy-two subjects took part in Experiment 1.

2.2. Results

Five items were lost by experimenter error; 6 subjects were replaced for ignoring instructions (2), for more than 25% error (1), or for more than 25% of response times more than 2 SD away from the grand mean (3). Remaining wrong or outlying responses were also discarded.

We first examined our hypothesis that priming would be robust throughout: it is. Two-way ANOVAs with Priming (primed vs. unprimed) and Word form (read vs. spontaneous) as crossed repeated measures showed a significant main effect of priming (MinF' (1,98) = 22.671, $p < .005$). As the leftmost bars in Figure 1 show, an effect of quality was also found. The read forms give a greater priming effect (94ms) than the more reduced spontaneous forms (44ms) (F_1 (1, 71) = 8.39, $p < .005$; F_2 (1, 55) = 4.83, $p = .032$ - all Neuman-Keuls tests were significant at $p < .05$ or better). Both priming effects are significant (read primed vs. unprimed tokens (t(70)= 7.786, $p < .001$; spontaneous primed vs. unprimed tokens (t (70) = 2.951, $p < .005$).

Next we compared the items for which the context that licenses assimilation/reduction would come **after** the token itself (nasals and d-deletions) with those which contain their own licensing environment (**within**) (weak initial and weak non-initial syllable word). A three-way by-materials ANOVA with Priming (primed vs. unprimed), Word form (read vs. spontaneous) and Licensing type (after vs. within) as repeated measures was used. Although the priming effect is robust and does not depend on the licensing type (Main effect of priming F_2 (1, 48) = 30.35, $p < 0.0001$; interaction priming by licensing $F_2 < 1$), the most effective primes were read tokens of the **within** group. These primed more effectively than read tokens of the **after** group or than spontaneous tokens from either group. (F_2 (1, 48) = 4.81, $p < 0.05$; all Neuman-Keuls tests were significant at $p < .05$ or better).

We see that even heavily assimilated, unintelligible tokens such as those typical of spontaneous speech can act as primes to facilitate access to mental lexical representations when they are

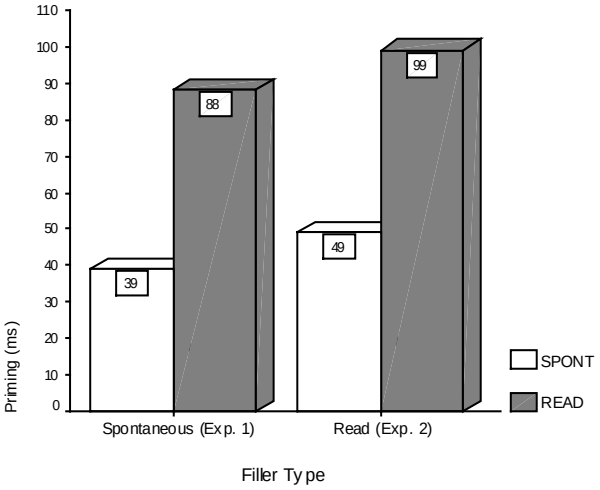

Figure 1: Priming effects for read and spontaneous primes in two filler environments spontaneous fillers (Exp. 1) and read fillers (Exp. 2).

presented with a local context which itself is assimilated. Recall that all of the fillers in Experiment 1 were themselves from spontaneous speech. But would running speech tokens serve as valid primes if the local context were altogether more canonical?

Experiment 2 addresses this issue by replacing all of the spontaneous filler tokens of words in Experiment 1 with read tokens of the same words read by the same speakers.

3. EXPERIMENT 2

3.1. Method

All filler words used in Experiment 1 were replaced by their read versions. Otherwise, materials and method were identical to those used in Experiment 1. Another 48 subjects from the population used in Experiment 1 were tested. Of these, 4 were discarded and replaced, 1 for not following instructions, another for a 25% error rate in all cells, and 2 others because their response times average more than 2s.d.'s above the group average in all cells of the design.

3.2. Results

The same analysis method was employed as for Experiment 1. Once again we found that priming was robust; (Main effect from priming F_1 (1, 47) = 26.21, $p < .0001$, F_2 (1, 52) = 30.91, $p < .0001$). Read forms were the more effective primes but the interaction was significant only on the by-subjects analysis. (priming x token form (F_1 (1, 47) = 6.72, $p < .02$; F_2 (1, 52) = 1.79, $p > .10$).

Next we repeated the analysis of licensing environment. A robust priming effect obtains (F_2(1, 51) = 35.07, $p < 0.0001$) but its interaction with type of prime was not significant.

3.3. Discussion

As figure 1 shows, the results of Experiment 2 did not differ substantially from those of Experiment 1: the use of clear filler items did not destroy the ability of spontaneous tokens to prime the visual counterparts. The major result of this manipulation appeared to be increasing underlying variance sufficiently to weaken interactions. A direct comparison is offered in Section 4.

4. COMPARISON OF EXPERIMENTS

The principal motivation for Experiment 2 was to examine the effect of ambient speech quality on priming. In comparing the outcomes of the two experiments we used priming as the dependent variable, that is, the difference between positive lexical decision times for matching and non-matching auditory-visual pairs. In the first ANOVA the factors were Ambient Clarity (Exp. 1 -spontaneous vs. Exp. 2 – read), Licensing type (after vs. within) and Word form (read vs. spontaneous). Faulty items from either experiment were removed in the comparison.

There was no main effect of ambient clarity on priming ($F_2 < 1.$) Across both experiments the read tokens primed more strongly than spontaneous tokens (Word form F_2 (1, 45) = 11.15, $p < 0.002$). Contrary to prediction this effect of quality was not altered by the nature of the fillers (Clarity x Word form: F_2 (1, 45) = 2.5, $p > .10$).

As Figure 2 shows, over both experiments words of the **within** group were more effective primes than those in the **after** group F_2 (1, 45) = 4.9, $p < 0.05$). Ambient clarity did not affect the results. ($F_{2<} < 1$).

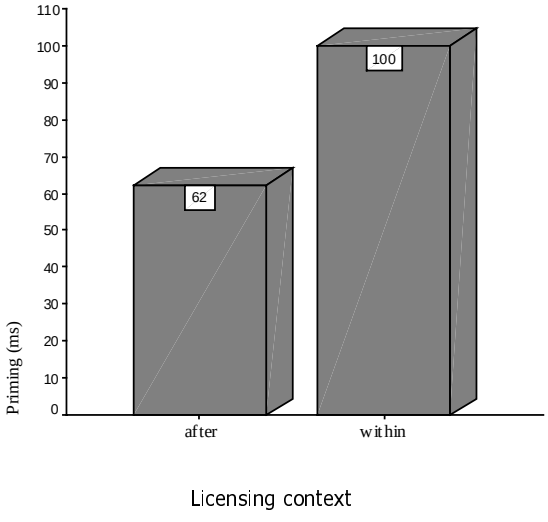

Figure 2: Difference in priming between tokens whose word-internal structure licenses reduction (**within**) and those licensed by following words (**after**).

5. CONCLUSIONS

The two experiments tested four hypotheses. The results bear on all of them. The quality hypothesis predicted that reduced tokens would prime more weakly than more canonical instances. This prediction was upheld in Experiment 1, but the effect was attenuated in Experiment 2. The robust access hypothesis predicted that nonetheless reduced tokens would allow lexical access. This effect also obtained. The ambient context hypothesis predicted weaker lexical access for more assimilated tokens when heard in lists of clear items. No such effects were found. Finally, the licensing hypothesis predicted more effective priming by assimilated or reduced tokens whose conditioning environments are word internal. Overall, such words proved to be the more effective primes.

The tokens used in the present experiments reflect naturally produced read and spontaneous speech more closely than those of previous studies which may have included more exaggerated assimilatory or reduced forms than is natural. The tokens used here exhibit various kinds of fast speech phenomena that naturally co-occur. Despite their poor quality and their presentation in isolation, they significantly facilitate lexical access. Accordingly, lexical access for running speech is likely to be more robust than recent experiments have suggested.

Nonetheless, distorted tokens impede lexical access to some degree. Our experiments instantiate a paradigm close to the techniques of Andruski et al. [4] and Connine et al. [3] who proposed that a goodness of fit metric operates during the recognition of distorted tokens. By comparing more and less assimilated tokens, our experiments suggest that normal lexical access is a task of variable difficulty.

It does not, however, appear that local speech quality plays a significant role. Our failure to find effects of filler quality suggests that the phonetic quality of a word affects lexical access for the word itself but has little further consequence for bottom-up processes operating on non-adjacent words. If lexical access is independent in this way, naturally occurring modulations of quality should not have catastrophic effects on comprehension of running speech.

Finally, the licensing effect suggests an extension of phonological inferencing [6], [10]. The presence of a licensing context appears to facilitate lexical access for assimilated/reduced tokens but not to be a necessary condition. Our results in this area are preliminary; they deserve to be pursued with more elaborately controlled stimulus sets. As they stand, the results are consistent with the notion of a licensing continuum rather than a system in which phonological changes are either licensed or blocked. Short stretches of speech, within words and around their boundaries, may be consulted to determine both which phonological events would allow the speech sounds to be mapped onto known words and how likely those events are.

6. REFERENCES

1. Marslen-Wilson, W.D. & Zwitserlood, P. “Accessing spoken words – the importance of word onsets,” *JEP:HPP* 15: 576-585, 1989.
2. Marslen-Wilson, W.D., Moss, H. & VanHalen, S. “Perceptual distance and competition in lexical access,” *JEP:HPP* 22: 1376-92, 1996.
3. Andruski, J., Blumstein, S. & Burton, M. “The effect of subphonetic differences on lexical access,” Cognition 52: 163-187, 1994.
4. Connine, C.M., Titone, D, Deelman, T. & Blasko, D. “Similarity mapping in spoken word recognition,” *Journal of Memory and Language* 37: 463-480, 1997.
5. Lahiri, A. & Marslen-Wilson, W. “The mental representation of lexical form,” *Cognition* 38: 254-294, 1991
6. Gaskell, G. & Marslen-Wilson, W. “Phonological variation and inference in lexical access,” *JEP:HPP* 22: 144-158, 1996.
7. Marslen-Wilson, W., Tyler, L.K., Waksler, R & Older L. "Morphology and meaning in the English mental lexicon *Psych. Rev*. 101: 3-33, 1994.
8. Anderson, A. et al. “The HCRC Map Task Corpus,” *Language and Speech* 34: 351-366, 1991.
9. Sotillo, C. “Phonological reduction and intelligibility in task-orientated dialogue,” *Unpublished Doctoral Dissertation, University of Edinburgh.*
10. Marslen-Wilson, W. & Gaskell, G. “Match and mismatch in lexical access,” *Paper presented to XXV Int. Congress of Psychology, Brussels*, 1992.