Markedness Effects in Visual Processing of Non-native Onset Clusters

Mackenzie Young and Colin Wilson

1. Introduction

A variety of experimental methods have provided evidence for knowledge of constraints on wordforms, including sensitivity to markedness distinctions among different types of unattested sequences. For example, in acceptability judgment tasks with both auditory and visual stimuli the rating (or binary judgment) of a form generally decreases with markedness (auditory stimuli: Scholes, 1966; Coleman & Pierrehumbert, 1997; Frisch, Large & Pisoni, 2000; Albright, 2009 [based on Albright & Hayes, 2003]; Kager & Pater, 2012; Hayes & White, 2015; visual stimuli: Bailey & Hahn, 2001; Daland et al., 2011). Many of the relevant studies have focused on markedness relations among onset clusters, for example the graded distinctions of well-formedness due to sonority sequencing (e.g., sonority fall < sonority plateau < small sonority rise < large sonority rise). Unsurprisingly, novel words beginning with onset clusters that are attested in the native language of the participants (e.g., *brif*) receive higher ratings and endorsement rates than otherwise matched nonwords beginning with unattested clusters (e.g., *rbif*). More interesting is the fact that judgments distinguish among unattested clusters in a way that reflects degrees of markedness. For example, nonwords beginning with a small-rise onset (e.g., *bnif*) tend to be judged as superior to nonwords beginning with falling-sonority onsets (e.g., *rbif*), even by speakers of languages in which neither onset type occurs in lexical forms.

Previous studies using auditory stimuli have also found sensitivity to markedness distinctions in online processing. More specifically, there is evidence from many perception and production tasks that nonnative clusters undergo the repair of perceptual epenthesis—the insertion of a reduced vowel that is not presented in the target form—at a way that reflects sonority-based and other markedness constraints (speech production: Davidson, 2006, 2010; Haunz, 2007; Wilson et al., 2014; speech perception: Dupoux et al, 1999; Berent et al., 2007; Maïonchi-Pino et al., 2014; Daland et al., submitted). For example, the results of syllable counting tasks essentially mirror those of acceptability judgment: monosyllabic nonwords beginning with more marked clusters are more likely to be identified as disyllabic, plausibly reflecting differences in the rate of perceptual epenthesis (e.g., epenthesis into a sonority plateau [bdif] → [bədif] applying more frequently than epenthesis into a small sonority rise [bnif] → [bənif]).

While perceptual epenthesis is robustly attested for auditory stimuli, only a small number of studies have reported parallel effects for visually-presented forms (Berent, 2008; Berent & Lennertz, 2010). If the same or similar repair process applies regardless of the modality in which stimuli are presented, this would support a more abstract phonological locus of perceptual epenthesis. On the other hand, failure to find evidence of this repair in visual processing would be consistent with other results indicating that high rates of perceptual epenthesis depend on the presence of particular acoustic cues such as stop releases (e.g., Davidson 2010; Wilson et al., 2014; Lennertz & Berent, 2015; Zhao & Berent, 2015; Daland et al., submitted).

In the present paper, we provide new evidence relevant to the question of whether marked onsets are repaired consistently across the auditory and visual modalities. Replicating the results of Berent & Lerentz (2010), we establish that sensitivity to sonority sequencing constraints can be found with

orthographic stimuli in a speeded same-different task involving onset clusters and their epenthetic counterparts (e.g., *rkip-REKIP*; Experiments 1-2). However, these initial findings are consistent with at least two hypotheses: orthographic stimuli beginning with marked onsets may be subject to epenthesis repair like their auditory counterparts; alternatively, more marked orthographic stimuli may be encoded less precisely in working memory, making the same-different comparison task more challenging. These hypotheses make distinct predictions about performance on comparison of stimuli that differ in ways other than the epenthesis repair (e.g., *rkip-RKIPE*). The results of Experiment 3 speak in favor of the encoding precision hypothesis, as parallel markedness effects are found for both repair and non-repair orthographic comparisons. In the final sections of the paper, we formalize the encoding hypothesis within a general framework of gradient symbolic representation and discuss the implications and some future extensions of our findings.

2. Experiment 1

The speeded same-different task has been widely used in the study of cognitive representations and processes (e.g., Ratcliff, 1981). On each trial, the participant is presented with a first stimulus (the *target*) which must be held in memory during a delay interval. Subsequent to the delay, a second stimulus (the *probe*) is presented and the participant responds as quickly and accurately as possible whether the probe is the 'same' as the target or different. In the orthographic version of the task, the target and probe are typically presented in different cases to prevent comparison on the basis of low-level visual features (and, for related reasons, a pattern mask is displayed during the delay period; see Figure 1). The intended notion of 'same' is

Same	Different Repair	Different Non-repair
rkip – RKIP	rkip – REKIP	rkip – RKIPE*
(rekip – REKIP)	(rekip – RKIP)	(rkipe – RKIP*)
(rkipe – RKIPE*)		

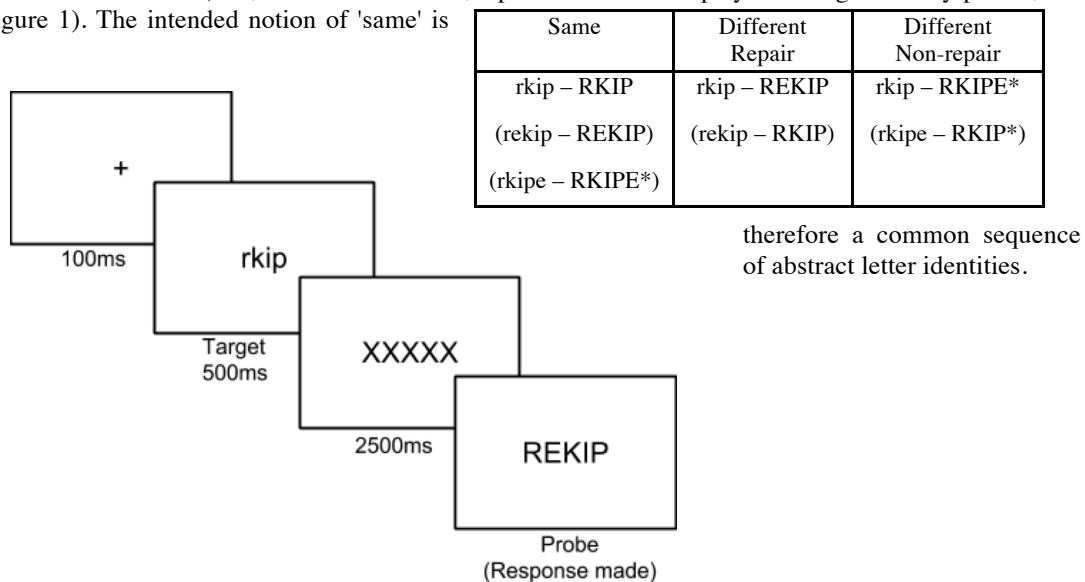therefore a common sequence of abstract letter identities.

Figure 1. Speeded orthographic same-different matching task and critical trial types (with fillers in parentheses). An asterisk indicates that a trial type appeared in Experiment 3 only.

In the first experiment, we attempted to directly replicate the previous findings of Berent & Lennertz (2010, Experiment 1) using exactly the same materials and procedure.

2.1 Participants

In each of the experiments reported here, twenty-four native English speakers at the Johns Hopkins University participated for partial course credit. No individual performed more than one experiment.

2.2 Materials

As noted above, the stimuli for this experiment were drawn from Berent & Lenertz (2010). In critical trials, the target was a CCVC nonword[1] that began with an attested large sonority rise (e.g., *bl*, *dr*), small sonority rise (e.g., **bn*, **dl*), sonority plateau (e.g., **pt*, **bd*), or sonority fall (e.g., **lb*, **rd*). The nucleus was always a single vowel letter, and the coda was either a single consonant letter or the digraph *th*. The probe was the same CCVC nonword (in upper case) or its epenthetic CECVC form (e.g., *rkip-RKIP* vs. *rkip-REKIP*; see Figure 1). Filler trials began with the corresponding CECVC forms and had cluster or identity probes (e.g., *rekip-RKIP* or *rekip-REKIP*). Assignment of items to the four critical and filler conditions was counterbalanced across participants, and item order was randomized separately for each participant.

2.3 Procedure

The same procedure was used for all three experiments reported here. Each trial began with the presentation of a lower-case target stimulus for 500 ms (e.g., *rkip*), followed by a pattern mask and memory interval (2500 ms), after which the probe stimulus appeared in upper case (e.g., *RKIP* or *REKIP*). Participants responded by pressing one button of a response box if the target and probe were perceived as the same sequence of letters (ignoring case), and another button if they were perceived as different letter sequences. Performance on this task was highly accurate, therefore our analyses focus on response time (RT) from the onset of the probe. There were 112 trials per participant in Experiment 1, and 200 trials per participant in Experiments 2 and 3.

2.4 Results

Responses that took longer than 1200 ms (which triggered a "too slow!" warning on the screen) were excluded (1.67%) from the analysis, as were trials in which the participant responded incorrectly (6.58%). A mixed-effects linear regression was conducted on response times. The fixed effects of identity (same vs. different) and sonority profile were sum-coded, and the random effect structure of the analysis was maximal. The results indicated a significant main effect of stimulus identity, with responses to same trials overall faster than responses to different trials ($\beta = -37$, $p < .01$; see Proctor et al. 1984). There was one marginal effect of the sonority profile of the initial cluster, indicating that responses were somewhat slower when the target began with a falling sonority cluster ($\beta = +13$, $p = .09$). While this experiment failed to replicate the finding of Berent & Lennertz (2010), in spite of the identity of materials and design, we wondered if more robust sonority effects would be revealed with slightly greater stimulus control.

3. Experiment 2
3.1 Materials

The stimuli in Experiment 2 were created by recombining the onsets and rimes of Experiment 1 in a way that controlled for orthographic neighborhood density across attested and unattested onsets. A word was considered an orthographic neighbor of a stimulus item if it could be made by adding, substituting, or deleting one letter in the item. Half of the items in this set had attested onsets and half had unattested onsets. Of the items with attested onsets, half had higher neighborhood density (ND = 5-7) and half had lower neighborhood density (ND = 0 or 1), and the same was true of the items with unattested onsets.

3.2 Results

As before, trials with response times greater than 1200 ms (1.44%) and trials with incorrect responses (7.67%) were removed prior to analysis. A mixed-effects linear regression with maximal random effect structure revealed that same trials again elicited faster responses than different trials (β

---

[1] While stimuli were generally nonwords, a few existing lexical items (e.g., *trap*, *drape*) were unintentionally included in each experiment.

= –27, $p < .01$). There were no significant differences among the sonority profiles, with the exception that responses were overall slower to falling-sonority items ($\beta = +17$, $p < .01$). This result, which agrees with previously reported findings, suggests that at least clusters with the most marked sonority profile are associated with some form of processing difficulty in the matching task.

3.3 Discussion

The results of this experiment are consistent with the hypothesis that, at some level of representation, orthographic stimuli are subject to an epenthesis repair in a way similar to auditory stimuli (Berent & Lennertz 2010, see Figure 2). In particular, subsequent to orthography-to-phonology conversion epenthesis could apply to the representation of a highly marked cluster as in *rkip* but not to that of a (phonologically) attested cluster as in *krip*. Slower processing for the former would result from the fact that, once the repair has applied, the phonological representation of the target matches that of the epenthetic probe. That is, 'different' responses to *rkip-REKIP* would be impeded by identity of the phonological representations of the two stimuli. (Note that according to this hypothesis no repair applies to the orthographic representation directly, only to its phonological recoding.)
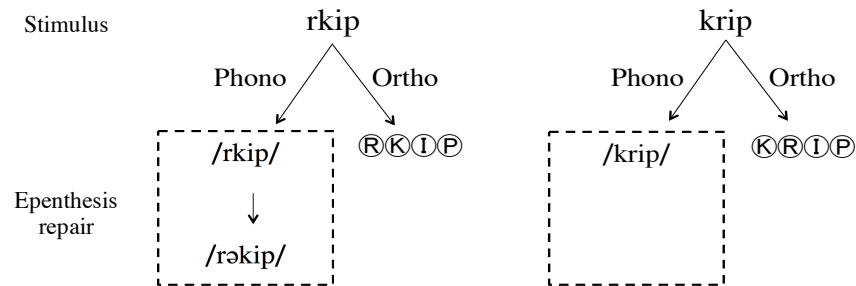

Figure 2. Repair hypothesis: After orthography-to-phonology conversion, the same epenthetic repair applies to the cluster as in auditory experiments. The likelihood of repair increases with cluster markedness.

However, the repair hypothesis is not the only possible explanation for the preceding results. According to the *encoding precision* hypothesis, the fidelity of the representation of an orthographic stimulus depends on its markedness: all other things being equal, more marked forms have less precise representations (see Figure 3). As in the repair hypothesis, a processing slow-down for nonwords beginning with ill-formed clusters would result from greater representational similarity between targets and probes, but similarity would not arise as the result of a repair process. Rather, imprecision itself would lead to partial matching of the encoding of the target and that of the probe.
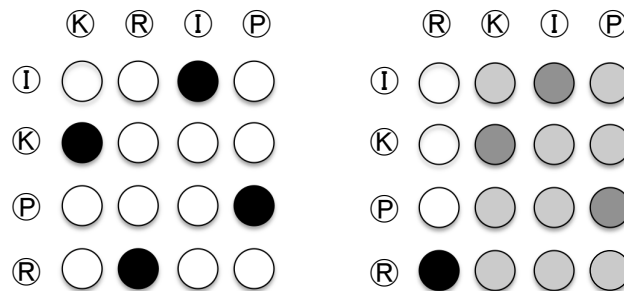

Figure 3. Encoding precision hypothesis: Stimuli with attested onsets are encoded with more precision than stimuli with unattested onsets. Letter position could also affect encoding strength; for example, the letter in the first position may be perceived with greater fidelity and therefore encoded more strongly for both stimulus types.

Previous experiments have been unable to distinguish between the repair and precision hypotheses, because non-identity probes have always corresponded to hypothesized repairs. In the

final experiment, we address this confound with a novel type of comparison.

4. Experiment 3
4.1 Stimuli

The critical target stimuli in Experiment 3 were the same as those of Experiment 2. In addition to the repair trials of Experiment 2 (e.g., *rkip-REKIP*), this experiment included an additional critical trial type in which epenthesis applied word-finally instead of within the cluster (e.g., *rkip-RKIPE*; see Figure 1). Fillers matched to the new critical trials were also included (e.g, *rkipe-RKIP*, *rkipe-RKIPE*).

4.2 Results and discussion

The responses from one participant, who responded 'different' on 87% of the trials, were excluded; therefore, the following analysis reflects the responses from the remaining 23 participants. As before, slow responses (1.89%) and errors (8%) were also excluded. A mixed-effects analysis of RTs established that participants were faster on same trials (β = -37, p < .01) and slower on trials in which the first stimulus began with a sonority fall (β = +13, p < .05). There were no further effects of sonority profile and no significant interaction between sonority and epenthesis location (i.e., repair CECVC vs. non-repair CCVCE; all ps > 0.65). Separate analyses of the CCVC-CECVC and CCVC-CCVCE critical pairs established that the effect of falling sonority was significant for both types.

This pattern of results obtained in this experiment are consistent with the encoding precision hypothesis, but unexpected from the perspective of the alternative repair hypothesis. Unlike insertion within the cluster, insertion at the end of the word does not repair the cluster. It is therefore unlikely that the phonological encoding of a form such as *rkip* would have a spurious word-final vowel or other property that could match the phonological encoding of the probe *RKIPE*. Indeed, as discussed more fully in section 6, the phonological representations of *rkip* and *RKIPE* are likely to be highly distinct rather than confusable. In the following section, we develop a formal version of the encoding precision hypothesis that does account for the slow-down in 'different' responses to such pairs.

5. Encoding precision model

The model combines ideas from previous research on subsymbolic computation (Smolensky & Legendre, 2006; Smolensky et al., 2010), the role of perceptual expertise in encoding and working memory (e.g., Curby et al., 2009; Lorenc et al., 2014), and the relationship between cognitive processes and information theory (e.g., Sims, 2016). The central claim of the model is that nonword targets containing more marked clusters are encoded in working memory with lower fidelity or *precision*. The effect on precision is not localized to the offending cluster, but instead impacts the encoding of many (if not all) letters of the target. Lower precision entails greater *similarity* between the encoding of the target and of many alternative letter strings, including both repair and non-repair probe types investigated in Experiment 3. To the extent that distinct strings have similar encodings, 'different' responses are predicted to be slower in the experimental task.

The subsymbolic component of our model involves embedding letter strings in a continuous vector space (e.g., Smolensky & Legendre, 2006: Chapter 5). Each letter identity (i.e., each letter abstracting away from case and other variations) is represented by a *filler* vector $\mathbf{v}_f$ and each letter string position is represented by a *role* vector $\mathbf{v}_r$. In the current implementation, we adopt a purely local and orthonormal encoding of fillers and roles. The *i*th filler is represented by a vector with *i*th component equal to 1 and 0s elsewhere (i.e., assuming alphabetic ordering, the letter identity Ⓐ corresponds to the vector [1 0 0 …]). In addition to the identities of all letters, we include a special *empty* filler that signals that no letter is present at a particular position. On the basis of evidence that letters are mentally coded relative to both edges of a string (Fischer-Baum et al., 2011), we employ a dual system of role vectors: one set of orthonormal vectors represents left-to-right serial positions, and another set (designated by '-') represents right-to-left positions. For convenience we restrict the length of letter strings that can be represented in the model, and hence the number of position roles in each set, to six. The vector representation of a letter string such as *rkip* is formed by identifying the

filler/role *bindings* that it contains (i.e., Ⓡ/1, Ⓚ/2, …, Ⓘ/-2, Ⓟ/-1) and adding together the *tensor products* of the corresponding filler and role vectors (i.e., $\mathbf{v}(rkip) = \mathbf{v}_Ⓡ \otimes \mathbf{v}_1 + \mathbf{v}_Ⓚ \otimes \mathbf{v}_2 + \cdots + \mathbf{v}_Ⓟ \otimes \mathbf{v}_{-2}$).

In addition to *pure* states such as the one just constructed, subsymbolic representation also admits *blend* states that represent weighted combinations of strings or other symbolic structures (Smolensky et al., 2010). In the current model, a blend arises whenever two or more fillers are partially present or activated in a single string position. One blend state, which we refer to as the *maximal blend* and denote simply by $\mathbf{v}$, is of particular importance here. The maximal blend has equal activation of all fillers in all roles; we assume in particular that all activations are equal to $1/|\Sigma|$ (where Σ is the alphabet of fillers). This state is the average of all letter strings that can be represented by the model. It expresses complete ignorance, or ambiguity, about the filler occupying each position—the maximal blend is the representation with the lowest possible precision.

As highlighted above, we assume that the precision of the encoding of a target is inversely related to its markedness. This assumption is formalized by taking a convex combination of the pure encoding of the target with the maximal blend state: for example, $\mathbf{e}(rkip) = \alpha \cdot \mathbf{v}(rkip) + (1-\alpha) \cdot \mathbf{v}$, where $\alpha \in [0,1]$. When α is near 1, the target is encoded with high precision: at each string position, all letter identities other than the correct one have negligible activations. As α approaches 0, the encoding of the target degenerates to the maximally ambiguous blend. Because our goal is only to account for qualitative effects on same-different performance, we make the minimal assumption that α lowers monotonically as cluster markedness increases (i.e., $\alpha_{\text{fall}} < \alpha_{\text{plateau}} < \alpha_{\text{small-rise}} < \alpha_{\text{large-rise}}$). Figure 3 schematizes the encoding of two targets differing in markedness, with grey shading indicating relative activation of a subset of fillers. As is common in the visual working memory literature, we assume that the probe (which is visible on-screen during the response period) is systematically encoded with maximal fidelity (here, $\alpha = 1$).

Given the memory encoding of the target, $\mathbf{e}$(target), and the encoding of a probe $\mathbf{e}$(probe), the model computes the cosine similarity (i.e., normalized dot product) of the two representations: $sim(\text{target, probe}) = \mathbf{e}(\text{target}) \cdot \mathbf{e}(\text{probe}) / (\|\mathbf{e}(\text{target})\| \ \|\mathbf{e}(\text{probe})\|)$. Across a range of α values (approximately $\alpha > 0.5$), the resulting similarities exhibit a clear pattern that accords with the response time effects observed experimentally. First, similarity is close to its maximum value of 1 whenever the target and probe are identical; this could underlie the 'fast same' effect observed across the experiments. Second, a target's similarity to the repair probe (e.g., $sim(rkip, REKIP)$) is provably identical to its similarity to the non-repair probe (e.g., $sim(rkip, RKIPE)$). This follows from the symmetry of the model's role representations, which do not privilege left-aligned over right-aligned string comparisons, together with the equal activation of all fillers in the maximal blend; it is consistent with the parallel effects of the two probe types in Experiment 3. As discussed earlier, the experimental parallelism is difficult to understand under the repair hypothesis—or indeed any account in which the effect of markedness is solely concentrated on the cluster itself rather than, as here, spread throughout the encoding of the target.

Finally, as anticipated, lower precision encodings (i.e., lower values of α) imply higher target-probe similarities. This reflects a trade-off between similarity due to positions in which the target and probe have the same letter identity and dissimilarity due to other positions. When α is near 1, only matching positions contribute substantially to similarity. With intermediate values of α (close to 0.5), matching positions contribute less but mismatching positions contribute more, because any letter identity partially matches the contents of the maximal blend. In more intuitive terms, lower precision encoding of the target produces a weaker 'different' signal to non-identical probes. If targets containing more marked clusters are encoded with lower precision, and if greater target-probe similarity leads to slower 'different' responses, as would be expected from process models of decision making (e.g., Ratcliff, 1981), markedness effects on same-different comparison are predicted across probe types.

Why are stimuli containing more marked clusters represented with lower fidelity? While we lack a complete answer to this question, we are encouraged by findings in other cognitive domains (e.g., face perception) that perceptual expertise leads to higher-precision representation of statistically expected or regular stimuli (e.g., upright faces) related to unexpected or irregular stimuli (e.g., upside down faces; Lorenc et al., 2014). More generally, if the representational capacity or resources of working memory are limited, as widely assumed, then representation of less probable stimuli is likely to require a loss of precision (e.g., Sims, 2016).

6. Discussion and conclusion

The results from these three experiments suggest that while structural constraints on clusters may be shared across modalities, their effects are more modality-specific. In speech, acoustic properties (e.g., the open transition between consonants) can account for specific repair effects for marked forms. In the absence of such acoustic properties (i.e., in reading), structural constraints affect the encoding strength of the stimulus, with is thus reflected as a processing slow-down. The differences in the pattern of results in auditory and visual experiments point to modality-specific perceptual challenges: speech perception requires segmentation of a continuous acoustic signal, which is subject to misinterpretation depending on acoustic cues and experience with certain sound sequences. Printed word recognition requires letter identity recognition and correct ordering of the letters, which is more difficult for certain letter positions due to crowding (see below).

In Experiments 2 and 3, we found evidence of processing difficulty for the nonwords beginning with the most marked clusters; only those nonwords beginning with falling sonority clusters consistently showed a response slow-down. The pattern of results across experiments suggests that orthographic stimuli are not subject to the same repair as auditory stimuli, but rather encoding strength varies with markedness. This conclusion is further supported by the results of Experiment 3, which showed the same pattern of results for cluster-internal repair trials (e.g., *rkip-REKIP*) as word-final non-repair trials (e.g., *rkip-RKIPE*). The repair hypothesis predicts that only repair trials should exhibit a response slow-down, since word-final epenthesis of *e* does not repair the cluster, and actually leads to greater phonological dissimilarity after encoding (e.g., *rkip*: /rəkip/ vs. *rkipe*: /rəkaɪp/). Again, only the encoding precision account predicts observed effects, and it makes the general prediction that knowledge of well-formedness should be reflected in speed of processing.

The encoding precision hypothesis requires further investigation, beginning with probing more directly participants' representations of briefly-presented attested and unattested onset clusters. One way to do so is with a full-report task, in which participants type in briefly presented nonwords. We conducted such an experiment using the same procedure as Experiments 1-3, with the only difference being that participants typed in the target rather than making a same-different judgment. The stimuli used in this experiment were the same items as Experiments 2 and 3. We found that participants were highly accurate, as expected, and did not consistently make repair errors when the target contained an unattested onset. The rate of errors that repaired the marked target in some way (17.32%) was almost the same as (and even lower than) the rate of errors that resulted in a more marked target (18.99%). The results from this study shed light on another possible explanation for the pattern of results from the speeded orthographic same-different studies. It could be argued that participants did not accurately represent the position of the epenthetic *e* in the repair and non-repair trials, since letter order is thought to be imprecise, which might have been the reason for the similar pattern of results across the two trial types. However, the results from the full-report experiment suggest that this is not the case.

A few additional studies focused on repair in marked orthographic stimuli may also benefit from the weak-encoding perspective. For example, using a lexical decision priming task in French, Sun & Peperkamp found evidence that nonword primes beginning with the marked onset *tl* (e.g., tlavier) facilitated lexical decisions on stimuli beginning with the onset *cl* (e.g., clavier) to a greater extent than nonword primes beginning with the attested onset *pl* (e.g., plavier). This result was taken as evidence for repair; since the cluster *tl* is often perceived as *cl* in auditory perception studies, phonological recoding of *tl* to *cl* may be the root of their priming effect. However, including an additional control could help to determine whether this result should be attributed to repair or weak encoding. It may be the case that nonwords beginning with marked clusters are better primes in this specific task, so including primes with another unattested onset, such as *rl* would be necessary to determine if the priming effect can truly be attributed to phonological repair of the prime. The weak-encoding hypothesis would predict that marked non-identity primes would generally facilitate lexical decision responses to a greater extent than non-identity primes with attested onsets, since weak-encoding of the marked prime leads to greater similarity to the target. Plausibly, the effects of imprecise encoding of marked orthographic structures could be detected with a variety of experimental methodologies.

References

Albright, Adam (2009). Feature-based generalisation as a source of gradient acceptability. *Phonology* 26.1: 9-41.

Bailey, Todd M., and Ulrike Hahn (2001). Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44.4: 568-591.

Berent, Iris, Donca Steriade, Tracy Lennertz, and Vered Vaknin (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition* 104.3: 591-630.

Berent, Iris (2008). Are phonological representations of printed and spoken language isomorphic? Evidence from the restrictions on unattested onsets. *Journal of Experimental Psychology: Human Perception and Performance* 34.5: 1288-1304.

Berent, Iris, and Tracy Lennertz (2010). Universal constraints on the sound structure of language: Phonological or acoustic? *Journal of Experimental Psychology: Human Perception and Performance* 36.1: 212-223.

Curby, Kim M., Kuba Glazek, and Isabel Gauthier (2009). A visual short-term memory advantage for objects of expertise. *Journal of Experimental Psychology: Human Perception and Performance* 35.1: 94-107.

Daland, Robert, Bruce Hayes, James White, Marc Garellek, Andrea Davis, and Ingrid Norrmann (2011). Explaining sonority projection effects. *Phonology* 28.2: 197-234.

Davidson, Lisa (2006). Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34.1: 104-137.

Davidson, Lisa (2010). Phonetic bases of similarities in cross-language production: Evidence from English and Catalan. *Journal of Phonetics* 38.2: 272-288.

Dupoux, Emmanuel, Kazuhiko Kakehi, Yuki Hirose, Christophe Pallier, and Jacques Mehler (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance* 25.6: 1568-1578.

Fischer-Baum, Simon, Jonathan Charny, and Michael McCloskey (2011). Both-edges representation of letter position in reading. *Psychonomic Bulletin & Review* 18.6: 1083-1089.

Frisch, Stefan A., Nathan R. Large, and David B. Pisoni (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language* 42.4: 481-496.

Haunz, Christine (2007). Factors in on-line loanword adaptation. PhD. dissertation, University of Edinburgh.

Hayes, Bruce, and James White (2015). Saltation and the P-map. *Phonology* 32.2: 267-302.

Kager, René, and Joe Pater (2012). Phonotactics as phonology: knowledge of a complex restriction in Dutch. *Phonology* 29.1: 81-111.

Lennertz, Tracy, and Iris Berent (2015). On the sonority levels of fricatives and stops. *The Mental Lexicon* 10.1: 88-132.

Lorenc, Elizabeth S., Michael S. Pratte, Christopher F. Angeloni, and Frank Tong (2014). Expertise for upright faces improves the precision but not the capacity of visual working memory. *Attention, Perception, & Psychophysics* 76.7: 1975-1984.

Maïonchi-Pino, Norbert, Bruno Cara, Jean Écalle, and Annie Magnan (2015). Is syllable segmentation developmentally constrained by consonant sonority within syllable boundaries in silent reading? Evidence in French children. *Journal of Research in Reading* 38.3: 226-248.

Proctor, Robert W., K. Venkata Rao, and Philip W. Hurst (1984). An examination of response bias in multiletter matching. *Perception & Psychophysics* 35.5: 464-476.

Ratcliff, Roger (1981). A theory of order relations in perceptual matching. *Psychological Review* 88.6: 552-572.

Sims, Chris R. (2016). Rate–distortion theory and human perception. *Cognition* 152: 181-198.

Smolensky, Paul, and Géraldine Legendre (2006). *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. MIT Press.

Smolensky, Paul, Matthew Goldrick, and Donald Mathis (2014). Optimization and quantization in gradient symbol systems: a framework for integrating the continuous and the discrete in cognition. *Cognitive Science* 38.6: 1102-1138.

Steriade, Donca (1997). Phonetics in phonology: the case of laryngeal neutralization. Ms., UCLA.

Sun, Yue, and Sharon Peperkamp (2015). The role of speech production in phonological decoding during visual word recognition: evidence from phonotactic repair. *Language, Cognition and Neuroscience* 31.3: 1-13.

Wilson, Colin, Lisa Davidson, and Sean Martin (2014). Effects of acoustic–phonetic detail on cross-language speech production. *Journal of Memory and Language* 77: 1-24.

Zhao, Xu, and Iris Berent (2015). Universal restrictions on syllable structure: Evidence from Mandarin Chinese. *Journal of Psycholinguistic Research*: 1-17.