

Predicting rhyme with segment confusion

1. Introduction

- 1.1 Some rhymes are better than other rhymes. Everyone would agree that *home – bone* is a better rhyme than *home – dose*.
- 1.2 Intuitively, *home – bone* is a better rhyme because those words are more similar to each other than the words in *home – dose* are.
- 1.3 Although there is general agreement that the goodness of a rhyme is determined by similarity, there are different suggestions for how rhyme similarity should be measured.
- 1.4 Some claim that the goodness of a rhyme is determined by similarity in articulatory features (Zwicky 1976, Berg 1990, Bauschatz 2003), others that it is determined by similarity in perception (Steriade 2003, Kawahara 2006).
- 1.5 Both sides point out individual cases that seem to argue against the other side (Zwicky 1976:689f, Berg 1990:22, Steriade 2003:593, Kawahara 2006:128ff.), but no full scale comparisons between the approaches exist.
- 1.6 The purpose of this study is just that – to provide a statistical comparison between the two approaches to see which is the better predictor of rhyme acceptability.
- 1.7 The results will show that similarity in perception is the better predictor.

2. Imperfect rhymes

- 2.1 Both the perceived similarity and the feature similarity between a segment and itself is 1 (Luce 1963:114).
- 2.2 The two approaches therefore make the same predictions for perfect rhymes, i.e. where the rhyme portion between two rhyming words is identical (*bone – stone*).
- 2.3 Just as the studies mentioned above, this study will therefore only focus on imperfect rhymes.

3. Gathering rhyme data

- 3.1 In order to study rhyme acceptability, we need a body of data on just that. That data does not exist.
- 3.2 All of the studies mentioned above use rhyme frequency data gathered from poetic corpora as a proxy for data on rhyme acceptability.
- 3.3 This choice is associated with three problems, addressed in the following.
- 3.4 To avoid these problems, we conducted an experiment where regular people gave their judgments on how acceptable a rhyme was.

4. Problem 1: Is acceptability correlated with probability?

- 4.1 An assumption in the studies above is that the acceptability of a rhyme determines how probable it is that a poet will employ that rhyme in his work (explicit only in Steriade 2003:583f.).
- 4.2 Intuitively, this seems right, and does not seem like a controversial assumption. However, there are no studies showing it is true (since there is no data on rhyme acceptability itself).
- 4.3 Gathering acceptability judgments directly avoids this issue.

5. Problem 2: Rhyming conventions

- 5.1 Another assumption in the studies above is that rhyme acceptability for poets is based on an unlearned and spontaneous measure of similarity, and not disturbed by rhyming conventions (Steriade 2003:584, 596, Kawahara 2006:133).
- 5.2 There are, however, several indications that conventions do play a significant role for imperfect rhymes in poetry.
- 5.3 Bauschatz (2003:33) finds that 83% of imperfect rhymes by English poets involve disagreement in fricative voicing, and he finds only occasional imperfect rhymes with nasal place disagreement. Hanson (2003:317) similarly notes that 90% of Pinsky's and 93% of Yeats' imperfect rhymes are fricatives disagreeing in voicing, and she finds only a few instances of nasal place disagreement. In rock lyrics, on the other hand, this pattern is reversed, with 49% of imperfect rhymes made up by examples of nasal place disagreement, and only occasional examples with fricative voicing disagreement (Zwicky 1976:685).
- 5.4 My interpretation of this finding is that 'professional' English poets are aware of the conventional acceptability of rhymes with fricative voicing disagreement, whereas the more 'naive' rock lyricists do not have this knowledge.
- 5.5 Bauschatz (2003:38) also finds that rhymes that were once perfect continue to form rhymes even when sound changes have made them imperfect. As an example, Shakespeare (16th century) rhymes *enter* with *venture* when both words had /t/, and Shelby (19th century) uses the exact same rhyme at a time when *venture* had changed its /t/ to /tʃ/.
- 5.6 A more extreme version of this convention is the 'eye rhyme', where words that were once perfect rhymes and spelled accordingly, such as Shakespeare's and Wyatt's (16th century) rhyme *wind* – *mind* /wʌɪnd – mʌɪnd/, continue to form rhymes even when their later pronunciations do not warrant it (Wyld 1923:10ff., Kökeritz 1953:31f., 218, Lagassé 2000:2401). Eye rhymes are also commonly found in Norwegian poetry, where past tense forms in *-et* /-ət/ rhyme with neuter definite nouns in *-et* /-ə/, but not with infinitives in *-e* /-ə/.
- 5.7 Zwicky (1976:683, 692) observes that many of the imperfect rhymes found in rock music are perfect rhymes in some variety of English, and suggests therefore that these are conventionalized rhymes, a view also taken by Kiparsky (in Hanson 2003:318).
- 5.8 A clear example of conventionalized imperfect rhymes is found in German poetry, where poets often rhyme unrounded vowels with rounded vowels. For instance, Goethe rhymes /ge:n/ with /ʃø:n/ in *Der Erlkönig* and /tsi:t/ with /bly:t/ and /tsvaifl/ with /tɔifl/ in *Faust*. Berg (1990:20) concludes that this is a rhyming convention adopted from German dialects where vowels do not contrast in rounding.
- 5.9 Because of the problem of rhyming conventions, Ohala finds poetry only 'potentially useful' for synchronic phonology. He points out, however, that the problem is the available data, not the method, noting that "poetic conventions regarding permissible rhymes [...] are usually conservative and once established tend to persist", whereas rhyme data obtained "from individuals [...] who have an imperfect knowledge of the traditional poetic conventions" would be highly useful (1986:8).

6. Problem 3: Lacking bad rhymes

- 6.1 As pointed out by Kawahara (2006:120), poetic corpora provide many examples of good and somewhat good imperfect rhymes, but hardly any or no examples of bad and somewhat bad imperfect rhymes.
- 6.2 Although both theories of similarity make clear predictions for bad rhymes, poetic corpora cannot provide the material to test them.
- 6.3 When asking individuals for their rhyme judgments, the rhymes can be distributed evenly across the spectrum of bad to good rhymes, and thus provide better material for a statistical analysis.

7. Experiment

- 7.1 We asked 72 American English speakers to rate the goodness of 77 imperfect rhymes on a 1-7 scale.
- 7.2 All rhymes were English monosyllabic nouns differing only in their coda consonant, and they were presented to the participants graphically.
- 7.3 The rhymes were constructed based on a random selection from pairs of segments.
- 7.4 As part of the course requirements for a phonology class, students performed the random selection, constructed the rhyme nouns, and gathered the judgment data from participants. These students were Lauren Eby Clemens, Ruthe Foushee, Laura Grestenberger, Adrienne Howard, Jenny Lee, Patrick Rich, David Sawicki, and Paul VanKoughnett.

8. Measuring similarity – features

- 8.1 Zwicky (1976:686) and Berg (1990:9) note that imperfect rhymes tend to differ in only one articulatory feature, and it is claimed that the acceptability of these rhymes is determined by how much they differ in such features (Zwicky 1976:677, Bauschatz 2003:30).
- 8.2 I take this to mean that this theory predicts that the more features a rhyme differs in, the less acceptable it is (cf. Kawahara 2006:116, 119).
- 8.3 In order to estimate the role of articulatory features, all rhyme pairs in the experiment were coded with the total number of articulatory features that their coda consonants did not have identical values for, using the feature values in Hayes 2009.
- 8.4 Every pair was also coded with the total number of acoustic features they differed in, using the feature values in Jakobson et al. 1952.

9. Measuring similarity – perception

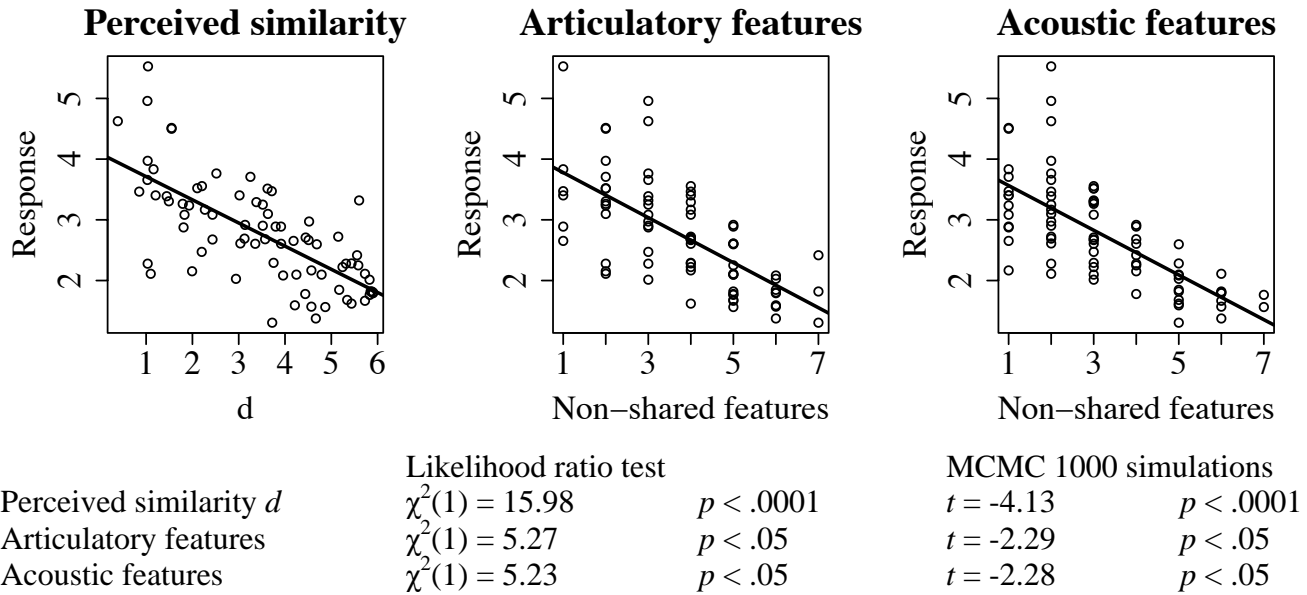
- 9.1 According to Steriade (2003:583f.), the acceptability of a rhyme is determined by how perceptually similar the words in the rhyme are. The more perceptually similar the words are, the more acceptable the rhyme is.
- 9.2 The perceptual similarity between two items is typically measured as a function of their confusability, meaning that measures of perceptual similarities require confusion matrices (Luce 1963:113, Macmillan & Creelman 2005:15).
- 9.3 Since confusion matrices do not exist for the languages investigated in Steriade 2003 and Kawahara 2006, the hypothesized correlation between perceptual similarity and rhyme acceptability could not be directly tested in those studies (cf. Kawahara 2006:132).
- 9.4 For American English, a complete confusion matrix for coda consonants in monosyllabic words exists in Cutler et al. 2004:3672.
- 9.5 Since Cutler et al. report the identification of coda consonants as proportions, we can use Luce's choice rule (1963:113f.) to calculate the perceptual similarities between these consonants, since this rule also refers to proportions.
- 9.6 According to Luce, the perceived similarity d between items x and y can be measured with the formula

$$d_{xy} = -\ln \left(\sqrt{\frac{p(r_y|s_x)p(r_x|s_y)}{p(r_y|s_y)p(r_x|s_x)}} \right)$$

- 9.7 Every rhyme pair in the experiment was therefore also coded with its perceptual similarity d value.

10. Model 1

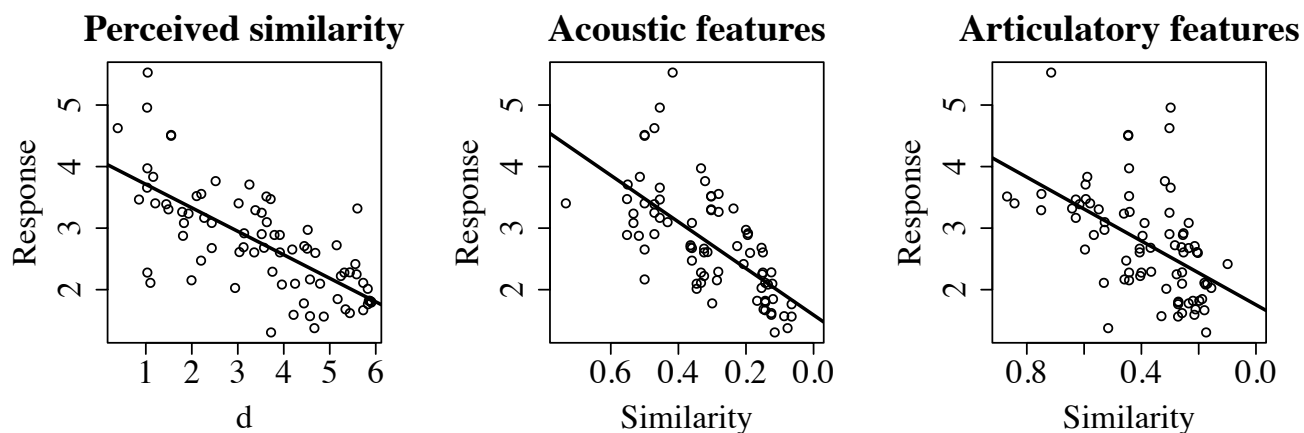
- 10.1 A linear mixed effects model was fitted to the data, with the participants' rating response as the dependent variable, and with participants and words as random effects.
- 10.2 The three measures of similarity were included as independent variables. These were:
- 10.2.1 Perceptual similarity d
 - 10.2.2 Number of non-shared articulatory features
 - 10.2.3 Number of non-shared acoustic features
- 10.3 The strength of these three measures can be seen from their χ^2 values from likelihood ratio tests and from their t -values in Markov Chain Monte Carlo sampling.
- 10.4 Both tests reveal that perceptual similarity is the best predictor of the participants' responses.



- Likelihood ratio tests measure the effect of dropping one similarity measure from a model with all three measures.
- MCMC simulation takes a model with all three similarity measures, and estimates the strength of one similarity measure residualized from the other two measures.

11. Model 2

- 11.1 Proponents of the feature similarity theory have also suggested that the patterns of imperfect rhymes could be better captured by referring to the natural classes of segments rather than treating each segment individually (Frawley 1984:46f., Bauschatz 2003:47ff.)
- 11.2 Frisch et al. (2004) similarly propose that similarity between segments is best captured by their natural classes, and devise a way of measuring similarity according to this principle.
- 11.3 In our second model, feature similarity was replaced by the similarity measure in Frisch et al. (2004) (Albright 2006). All other aspects of the model remain the same as in model 1.
- 11.4 Also in this case, perceptual similarity is the best predictor.



	Likelihood ratio test	MCMC 1000 simulations
Perceived similarity d	$\chi^2(1) = 19.18$ $p < .0001$	$t = -4.57$ $p < .0001$
Acoustic features	$\chi^2(1) = 7.83$ $p < .01$	$t = 2.81$ $p < .01$
Articulatory features	$\chi^2(1) = 3.31$ $p = .07$	$t = 1.8$ $p = .07$

12. Conclusion

- 12.1 With the use of rhyme acceptability data and a confusion matrix for coda consonants, it has been possible for the first time to fully compare the feature similarity theory with the perceptual similarity theory.
- 12.2 The results show that perceptual similarity is the best predictor of rhyme acceptability.
- 12.3 More than anything else, the acceptability of a rhyme is determined by how perceptually similar the rhyming items are, as suggested by Steriade (2003:583f.).

13. Discussion

- 13.1 The rich system of articulatory features does not suffice to account for the acceptability of rhymes, and does in fact do worse than the simple 12-feature system of acoustic features, as seen in model 2.
- 13.2 Perceptual similarity is the best predictor of rhyme acceptability, and can also adequately explain a range of phenomena in synchronic phonology (Kawahara 2006:114)
- 13.3 This does not mean, however, that the relation between phonological segments should be represented exclusively as perceptual distances, since some segmental relations are clearly better accounted for with the use of articulatory features (cf. Cristià & Seidl 2008).
- 13.4 If our model of phonological representations is meant to cover the wide range of attested synchronic patterns of segments, then a system is needed where reference is made both to articulation and perception (cf. Flemming 2002).

References

- Albright, Adam. 2006. Segmental Similarity Calculator. <http://web.mit.edu/albright/www/>
- Bauschatz, Paul. 2003. Rhyme and the Structure of English Consonants. *English Language and Linguistics* 7:29-56.
- Berg, Thomas. 1990. Unreine Reime als Evidenz für die Organisation phonologischer Merkmale. *Zeitschrift für Sprachwissenschaft* 9:3-27.
- Cristià, Alejandrina and Amanda Seidl. 2008. Is Infants' Learning of Sound Patterns Constrained by Phonological Features? *Language Learning and Development* 4:203-227.
- Cutler, Anne, Andrea Weber, Roel Smits, and Nicole Cooper. 2004. Pattern of English Phoneme Confusions by Native and Non-Native Listeners. *The Journal of the Acoustical Society of America* 116:3668-3678.

- Flemming, Edward Stanton. 2002. *Auditory Representations in Phonology*. New York, NY: Routledge.
- Frawley, William. 1984. A Note on the Phonology of Slant Rhyme. *Language and Style* 17:44-47.
- Frisch, Stefan A., Janet Breckenridge Pierrehumbert, and Michael B. Broe. 2004. Similarity Avoidance and the OCP. *Natural Language and Linguistics Theory* 22:179-228.
- Hanson, Kristin. 2003. Formal Variation in the Rhymes of Robert Pinsky's *The Inferno of Dante*. *Language and Literature* 12:309-337.
- Hayes, Bruce. 2009. *Introductory Phonology*. Oxford: Wiley-Blackwell.
- Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle. 1952. *Preliminaries to Speech Analysis. The Distinctive Features and Their Correlates*. Cambridge, MA: The MIT Press.
- Kawahara, Shigeto. 2006. Half Rhymes in Japanese Rap Lyrics and Knowledge of Similarity. *Journal of East Asian Linguistics* 16:113-144.
- Kökeritz, Helge. *Shakespeare's Pronunciation*. New Haven, CT: Yale University Press.
- Lagassé, Paul (ed.). 2000. *The Columbia Encyclopedia*. Columbia University Press.
- Luce, Robert Duncan. 1963. Detection and Recognition. In R. D. Luce, R. R. Bush, and E. Galanter, eds., *Handbook of Mathematical Psychology. Volume I, Chapters 1-8*, 103-189, New York, NY: John Wiley and Sons, Inc.
- Macmillan, Neil A. and C. Douglas Creelman. 2005. *Detection Theory: A User's Guide*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Ohala, John J. 1986. Consumer's Guide to Evidence in Phonology. *Phonology Yearbook* 3:3-26.
- Steriade, Donca. 2003. Knowledge of Similarity and Narrow Lexical Override. In P. M. Nowak, C. Yoquelet, and D. Mortensen, eds., *Proceedings of the Twenty-Ninth Annual Meeting of the Berkeley Linguistics Society. February 14-17, 2003. General Session and Parasession on Phonetic Sources of Phonological Patterns: Synchronic and Diachronic Explanations*, 583-598, Berkeley, CA: Berkeley Linguistics Society.
- Wyld, Henry Cecil. 1923. *Studies in English Rhymes from Surrey to Pope. A Chapter in the History of English*. London: John Murray.
- Zwicky, Arnold M. 1976. Well, This Rock and Roll Has Got to Stop. Junior's Head Is Hard as a Rock. In S. Mufwene, C. A. Walker, and S. B. Steever, eds., *Papers from the Twelfth Regional Meeting. Chicago Linguistic Society. April 23-25, 1976*, 676-697, Chicago, IL: Chicago Linguistic Society.