Confusion patterns and response bias in spoken word recognition of German disyllabic words and nonwords

2007 International Congress of the Phonetic Sciences, Saarbrücken, Germany

Speech Research Laboratory at Indiana University

Robert Felty

Research Goals

- Test the prediction made by combinatorial models of Lexical Access (e.g. Clahsen et al., 2001; Taft & Forster, 1975; Taft, 1988) that morphological complexity can affect language comprehension
- Compare effects of lexical status and neighborhood density to previous results from speech-in-noise tasks using CVC stimuli
- Determine how lexical and/or morphological information can bias perception of highly confusable phonemes

Method

Materials

- 150 nonwords and 150 German words (half monomorphemic and half bimorphemic) were selected from the CELEX database (Baayen & Rijn, 1993). Original analysis found that items containing post-vocalic /ʀ/ were treated quite differently than other stimuli, and therefore have been excluded from the analysis, leaving 94 nonwords, 36 mono- and 43 bimorphemic words for analysis.
- All stimuli were of the form CVCCVC (where V includes short and long vowels as well as diphthongs), with stress on the first syllable.
- Nonword stimuli were based upon the word stimuli such that the two sets were fairly phonemically balanced.

Participants

- 30 native speakers of German were recruited from the University of Konstanz (Germany)

Task — Speech-in-noise

- Participants heard the recorded materials over headphones and entered responses via keyboard.
- Signal dependent noise was added to the stimuli (Schroeder, 1968).

Analysis

- The data were analyzed using the j-factor model of Boothroyd & Nittrouer (1988).
- The j-factor model provides a measure of the number of independent units in a stimulus.
- A result of $j = n$ for **nonwords** (where n is the number of phonemes in the stimulus) can be interpreted as evidence that phonemes are perceived **independently** of each other.
- A result of $j < n$ for **words** has been interpreted as a bias towards words (Nearey, 2001).
- Previous studies using English stimuli have consistently found $j = 3$ for CVC nonwords, and $j \approx 2.5$ for CVC words (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003).

Predictions

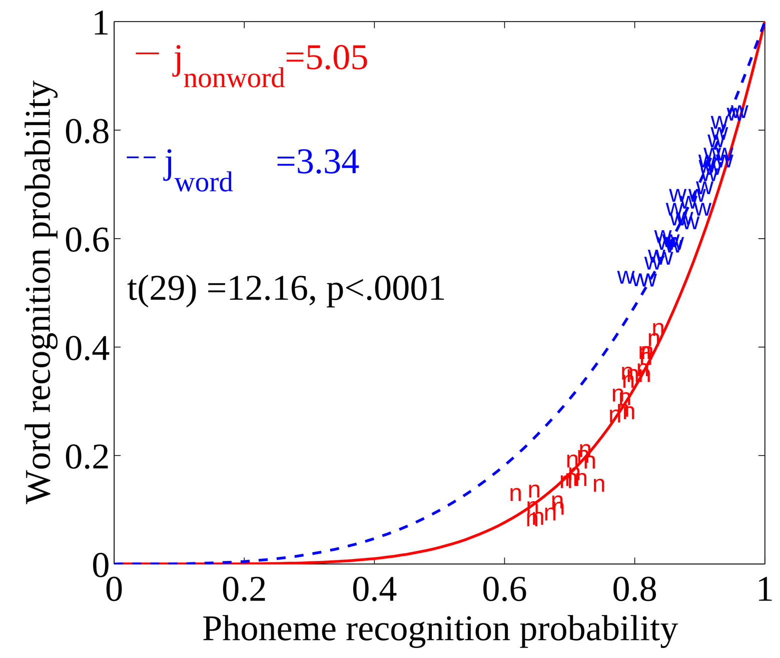
Although no studies to date have used the j-factor model to analyze disyllabic stimuli, several predictions can be made based on previous studies using the j-factor model with CVC stimuli (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003).

1. $j_{nonword} \approx 6$. This result would provide evidence that phonemes in nonwords are perceived independently of one another.
2. $j_{word} \approx 5$. Given that previous studies using CVC stimuli have found $j_{word} \approx 2.5$ (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003), it is logical to hypothesize that j_{word} will be twice as large for CVCCVC stimuli.
3. $j_{bi} > j_{mono}$. This prediction follows from the hypothesis that morphological units are stored in the lexicon, and that increasing the number of morphemes in a word should add to the number of independent units.
4. $j_{word} \propto$ density: Neighborhood density provides an inhibitory effect, such that words in dense neighborhoods are more difficult to process than words in sparse neighborhoods (Benkí, 2003).
5. Listeners rely more heavily on lexical and grammatical information in the absence of clear acoustic information. Thus, effects of morphology should be greatest for highly confusable stimuli.

J-Factor Analysis Results

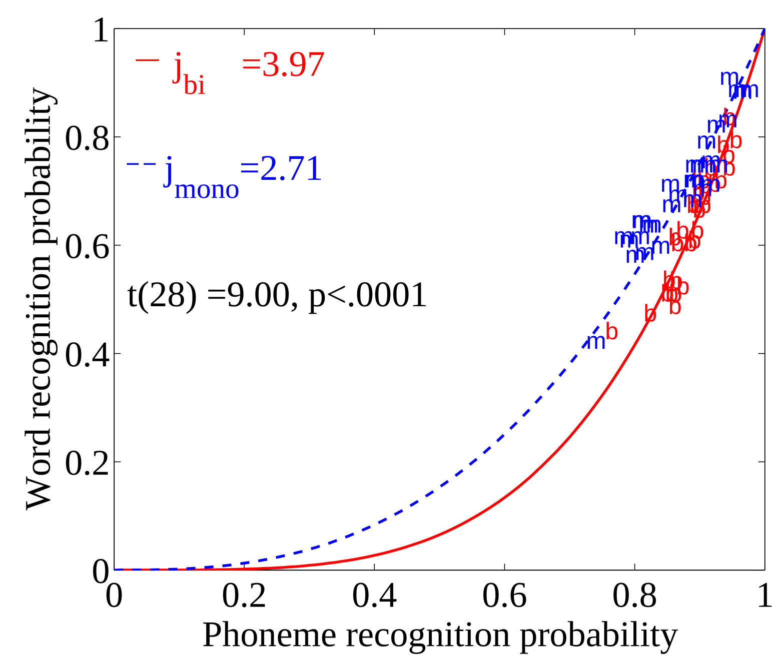
- The following figures display j-factor results for each of the context effects listed in the predictions.
- Curves represent $p_w = p_p^j$.
- Statistics shown in the figures are from paired t-tests; before computing the statistics, all points lying in the floor or ceiling ranges (> .95 or < .05) were removed, but are still shown on the plot.

Lexical Status



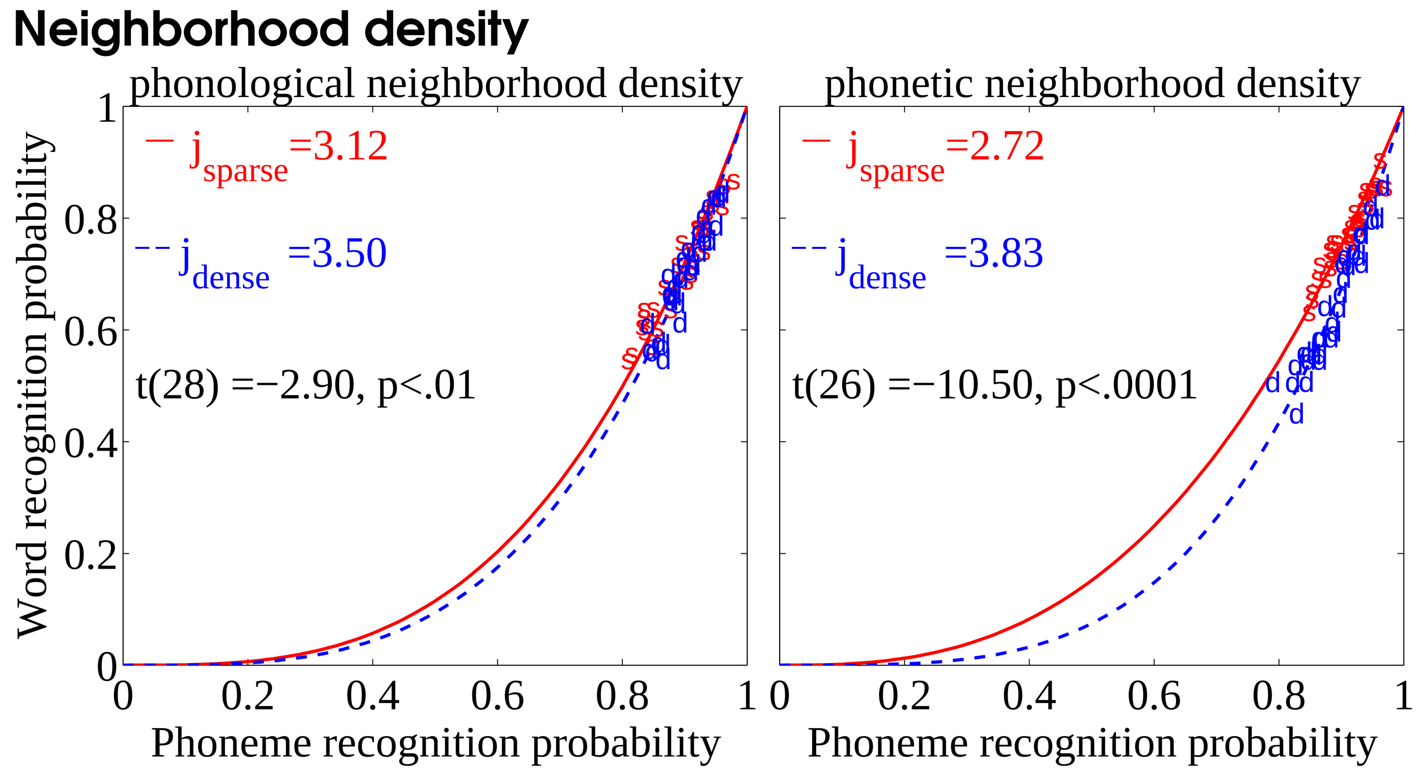
- As predicted, and consistent with previous studies (Boothroyd & Nittrouer, 1988; Olsen et al., 1997; Benkí, 2003) words had significantly lower j-scores than nonwords, indicating a facilitatory effect of lexical status.

Morphology



- As predicted, bimorphemic words had significantly higher j-scores than monomorphemic words.
- This can be seen as evidence that bimorphemic words are composed of more independent units than monomorphemic words, and that morphology can play a role in spoken word recognition.

Neighborhood density



- Two different measures of neighborhood density were calculated — a phonological measure, in which all words with an edit distance of 1 are treated as neighbors, e.g. *pat* has neighbors *pet* and *rat* (see Newman et al., 1997), and a phonetic measure was also calculated, based on the confusion matrices from the nonword data (see Luce & Pisoni, 1998; Benkí, 2003). The phonetic measure treats *pet* as a closer neighbor to *pat* than *rat*, given that /æ/ and /ɛ/ are more highly confusable than /ɹ/ and /p/.
- Sparse and dense groups were created from the word list using a median split.
- As predicted, and consistent with previous results (Benkí, 2003; Benkí & Felty, 2006), words in dense neighborhoods had significantly higher j-scores than words in sparse neighborhoods, indicating an inhibitory effect of neighborhood density.
  - The difference in j between sparse and dense neighborhoods is greater using the phonetic measure than the phonological measure.
  - Subsequent linear regression analyses showed that the phonetic measure accounted for 14.5% of the variation in j ($F(1, 182) = 13.78, p < .001$), while the regression analysis using the phonological measure was not significant.
  - These results are consistent with previous results using the j-factor model (Benkí, 2003; Benkí & Felty, 2006), and underscore the importance of including fine phonetic detail in models of spoken word recognition.

Perceptual distinctiveness, morphology, and response bias

- German inflectional suffixes can be used to investigate interactions between perceptual distinctiveness and response bias due to grammatical and/or statistical properties, as laid out in prediction 5.
- Of the inflectional suffixes in German, *-m* and *-n* are highly confusable, yet the *-n* ending occurs much more frequently.
- In order to investigate a possible interaction between morphology and response bias, a Signal Detection Theory (SDT) analysis was carried out.
- To carry out the SDT analysis, the original confusion matrices for each SNR were transformed into 2x2 submatrices. An SDT analysis was then applied to each submatrix.
- The table below displays the results of the SDT analysis for nonwords, mono- and bimorphemic words.
  - /m/ is the target stimulus.
  - Positive values of c indicate a bias towards /n/.
  - The final two columns list the total number of presentations of /m/ and /n/

SDT analysis of /m/ and /n/ confusions

	d'	c	/m/	/n/
Nonwords				
lower SNR (2 dB)	-0.182	0.555	240	240
higher SNR (7 dB)	0.664	0.743	240	240
Bimorphemes				
lower SNR (2 dB)	1.616	0.984	128	352
higher SNR (7 dB)	1.913	0.556	128	352
Monomorphemes				
lower SNR (2 dB)	3.514	0.239	48	192
higher SNR (7 dB)	4.733	-0.060	48	192

- In the absence of lexical context (nonword condition), /m/ and /n/ are highly confusable, with a small bias towards /n/.
- /m/ and /n/ are perceived as most distinct in the monomorphemic condition.
- Bias towards /n/ is greatest in the bimorphemic case.
- The SDT analysis lends greater support to the notion that morphology is encoded in the mental lexicon.

Conclusions

- The j-factor analysis showed that phonemes are perceived roughly independently of one another in nonwords, and that there is a strong bias towards words over nonwords.
- The difference in j between mono- and bimorphemic words suggests that morphological structure is encoded in the lexicon.
- Neighborhood density had a robust effect on word recognition, such that words in sparse neighborhoods showed a strong bias over words in dense neighborhoods. Moreover, a phonetically based measure of neighborhood density accounted for a much larger portion of the variation in the data than a phonologically based measure.
- Finally, an SDT analysis showed that listeners exploit statistical properties of the lexicon when faced with highly confusable phonemes.

References

Baayen, R., H & Rijn, H. (1993). The CELEX lexical database (cd-rom). Philadelphia: Linguistics Data Consortium, University of Pennsylvania.

Benkí, J. (2003). Quantitative evaluation of lexical status, word frequency and neighborhood density as context effects in spoken word recognition. *Journal of the Acoustical Society of America*, 113(3), 1689–1705.

Benkí, J. & Felty, R. (2006). Relative contributions of initial and final similarity to neighborhood density effects on english spoken word recognition. Poster presented at the Fifth International Conference on the Mental Lexicon: Montreal, Quebec.

Boothroyd, A. & Nittrouer, S. (1988). Mathematical treatment of context effects in phoneme and word recognition. *Journal of the Acoustical Society of America*, 84, 101–114.

Clahsen, H., Isenbeiss, S., Hadler, M., & Sonnenstuhl, I. (2001). The mental representations of inflected words: an experimental study of adjectives and verbs in German. *Language*, 77(3), 510–543.

Luce, P. & Pisoni, D. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, 19, 1–36.

Nearey, T. M. (2001). Phoneme-like units and speech perception. *Language and Cognitive Processes*, 16, 673–681.

Newman, R. S., Sawusch, J. R., & Luce, P. A. (1997). Lexical neighborhood effects in phonetic processing. *Journal of Experimental Psychology*, 23(1), 873–889.

Olsen, W., Tasell, D. V., & Speaks, C. (1997). Phoneme and word recognition for words in isolation and sentences. *Ear and Hearing*, 18(3), 175–188.

Schroeder, M. (1968). Reference signal for signal quality studies. *Journal of the Acoustical Society of America*, 44, 1735–1736.

Taft, M. (1988). A morphological decomposition model of lexical representation. *Linguistics*, 26, 657–667.

Taft, M. & Forster, K. (1975). Lexical storage and retrieval of prefixed words. *Journal of Verbal Learning and Verbal Behavior*, 14, 638–647.

Acknowledgements

I wish to thank the Linguistics department at the University of Konstanz for their hospitality and assistance, particularly Aditi Lahiri and Mathias Scharinger. I am also grateful for the support and insights of my advisors at the University of Michigan, José Benkí and Pam Beddor.