

RESEARCH ON SPOKEN LANGUAGE PROCESSING
Progress Report No. 22 (1998)
Indiana University

**Lexical Neighborhoods and Subjective Intelligibility Ratings:
A Preliminary Report¹**

Brian F. Bowdle and Richard Wright²

*Speech Research Laboratory
Department of Psychology
Indiana University
Bloomington, Indiana 47405*

¹ This work is support by NIH-NIDCD Research Grant DC00111 to Indiana University.

² Now at the Department of Linguistics, University of Washington, Seattle, WA.

Lexical Neighborhoods and Subjective Intelligibility Ratings: A Preliminary Report

Abstract. Recent research has shown that lexical properties known to influence word recognition – namely, frequency, neighborhood density, and neighborhood frequency – also affect the pronunciation of words (e.g., Balota & Chumbley, 1985; Goldinger & Summers, 1989; Wright, 1997). These findings could be explained within a cooperative framework, according to which speakers continually assess the ability of listeners to recover the message, and modify their pronunciations accordingly. However, this account assumes that speakers can somehow determine the effects of certain lexical properties on intelligibility. In this paper, we report a preliminary study in which subjects rated how intelligible they thought different words would be in an imagined noisy environment. The results suggest that speakers can indeed predict how lexical neighborhood characteristics influence word recognition.

Introduction

It is well known that speech production is influenced by both physical environment and linguistic context. At the physical level, speakers produce more careful and intelligible utterances in the presence of loud background noise (e.g., Lane & Tranel, 1971; Lane, Tranel, & Sisson, 1970), as well as when the listener is unable to see the speaker's face (Anderson, Bard, Sotillo, Newlands, & Doherty-Sneddon, 1997). At the linguistic level, words are pronounced with greater duration and clarity when they introduce new information into a discourse than when they refer to given information (e.g., Bolinger, 1963; Chafe, 1974; Hawkins & Warren, 1994). Likewise, the articulation of words tends to be reduced in sentential contexts where the words are highly predictable (e.g., Fowler & Housum, 1987; Hunnicutt, 1985; Lieberman, 1963).

These effects have typically been explained within a *cooperative framework*, according to which speakers continually assess the ability of listeners to recover the message, and modify their pronunciations accordingly (e.g., Bolinger, 1963; Chafe, 1974; Lieberman, 1963; Lindblom, 1990). On this view, speakers are aware of physical and linguistic factors that may interfere with comprehension, and produce more hyper-articulated speech in the presence of such factors to meet the needs of the listener.

However, physical environment and linguistic context are not the only factors that influence speech production. Recent research has shown that certain lexical properties may also affect the pronunciation of words. For example, Balota and Chumbley (1985) found that words with low usage frequencies are pronounced with greater duration than words with high usage frequencies. Likewise, Goldinger and Summers (1989) found that word pronunciation is influenced by *neighborhood density*, or the number of words that are phonologically similar to the target word. In their study, subjects read minimal pairs of English CVC words in which the word initial stop consonant was either voiced or voiceless. The voicing contrast (as measured by changes in voice onset time) was more extreme in pairs from dense neighborhoods than in pairs from sparse neighborhoods. Finally, Wright (1997) examined the pronunciation of words that varied in terms of three lexical properties: frequency, neighborhood density, and *neighborhood frequency*, or the mean usage frequency of the target word's phonological neighbors. Wright found that low frequency words from dense, high frequency neighborhoods showed a greater degree of vowel dispersion than high frequency words from sparse, low frequency neighborhoods.

What makes these findings particularly interesting is the fact that the lexical properties manipulated in the above production tasks also affects the recognition of words. Specifically, words that have low frequencies, high neighborhood densities, and high neighborhood frequencies are recognized slower and less accurately than words with the opposite characteristics (e.g., Goldinger, Luce, & Pisoni, 1989; Luce & Pisoni, 1998; Pisoni, Nusbaum, Luce, & Slowiaczek, 1985). Thus, those words that receive the greatest degree of hyper-articulation in production tasks are also found to be the least intelligible in recognition tasks. This suggests that the cooperative framework could be extended to explain the effects of lexical neighborhood characteristics on word pronunciation. On this view, speakers (1) have access to lexical properties such as frequency and neighborhood density, (2) recognize how these properties affect word recognition, and (3) can modify their pronunciations accordingly.

There is evidence that speakers can estimate the relative frequencies of words with reasonable accuracy (e.g., Carroll, 1971; Gernsbacher, 1984; Shapiro, 1969). However, it is not clear whether speakers can estimate neighborhood densities or neighborhood frequencies, let alone infer the relationship between any of these lexical properties and intelligibility. Indeed, it has been argued that the effects of lexical neighborhood characteristics on word pronunciation might more easily be explained at the level of motor programs (Balota, Boland, & Shields, 1989; Goldinger & Summers, 1989). For example, low frequency words might simply be associated with less familiar motor programs, which would lead to longer productions. Likewise, greater constraints may be imposed on the motor programs of words from dense neighborhoods to avoid productions of unintended, phonologically similar words, which would lead to more careful articulation.

Because the motor program account makes no assumptions about the speaker's knowledge of the relationship between lexical neighborhood characteristics and intelligibility, this explanation seems more parsimonious than that offered by the cooperative framework. However, there is little empirical evidence that directly favors one of these approaches over the other. The purpose of the present study was to test a key prediction of the cooperative approach – namely, that speakers can accurately assess the ability of listeners to recognize words with different lexical properties.

Method

Subjects

Twenty Indiana University undergraduates participated in partial fulfillment of a course requirement. All subjects were native speakers of English, and none reported having ever had a hearing or speech disorder.

Materials

The stimuli were taken from the Easy-Hard Multi-Talker Speech Database (Torretta, 1995). The database contains digital sound files of 150 monosyllabic English CVC words spoken by 10 talkers (five male and five female) at three speaking rates (slow, medium, and fast). Seventy-five of the words were labeled “easy” coming from sparse similarity neighborhoods and having high usage frequencies relative to their neighbors. The remaining 75 words were labeled “hard” coming from dense similarity neighborhoods and having low usage frequencies relative to their neighbors. Despite these lexical differences, all “easy” and “hard” words were highly familiar, scoring between 6.25 and 7.00 on the seven-point Hoosier Mental Lexicon scale (Nusbaum, Pisoni, & Davis, 1984).

For this study, only recordings of the first female talker (F1) at the medium-speaking rate were used. This set of sound files was selected for two reasons. First, the “easy” and “hard” words from this

set were equally high in intelligibility (Torretta, 1995). Second, the “easy” and “hard” words from this set showed the smallest difference in vowel dispersion (Wright, 1997). Thus, any difference in the subjective intelligibility of the “easy” and “hard” words should not be attributable to differences in their presentation level or acoustic characteristics.

Procedure

Subjects were run in groups of one to three in a quiet listening room. Each subject was seated in a separate booth in front of a computer monitor and a button box. Prior to running, subjects were told that they would hear a series of words, and that for each word they should “imagine the word being spoken in a noisy environment (such as a busy street corner or a crowded room), and estimate how intelligible the word would be for a listener standing a few feet away.” Subjective intelligibility ratings were made by pressing one of seven buttons on the button box. Each button corresponded to a value on a seven-point scale of intelligibility, with button 1 indicating that the word would be “very unintelligible” and button 7 indicating that the word would be “very intelligible.” Subjects were told that the more likely a listener would be to mishear a word or fail to catch it altogether, the lower the intelligibility rating should be.

Each subject received all 150 “easy” and “hard” words in a random order. The words were presented over matched and calibrated Beyerdynamic DT-100 headphones at a comfortable listening level (75 dB/SPL). Prior to the presentation of each word, a prompt appeared on the computer monitor. Once a response had been entered on the button box, the prompt appeared again and the next word was presented. Presentation of stimuli and response collection was controlled by a 200 MHz Gateway 2000 Pentium computer. Upon completion of the ratings task, subjects were given a post-test questionnaire in which they were asked to describe any strategies they had used to estimate how intelligible different words would be in a noisy environment.

Results and Discussion

Subjective Intelligibility Ratings

The subjective intelligibility of “easy” words ($M = 4.31$, $SD = 0.48$) was higher than that of “hard” words ($M = 3.54$, $SD = 0.55$). This difference was significant both by subjects ($t_s(19) = 7.37$, $p < .001$) and by items ($t_i(148) = 6.90$, $p < .001$). The complete distribution of the 75 “easy” and 75 “hard” words along the subjective intelligibility scale is shown in Figure 1.

 Insert Figure 1 about here

Subjects in this study were consistently able to determine that words from sparse neighborhoods and having high relative usage frequencies should be more easily identified than words from dense neighborhoods and having low relative usage frequencies. Indeed, 19 of the 20 subjects gave the “easy” words higher subjective intelligibility ratings than the “hard” words. The possession of such metalinguistic knowledge is consistent with the cooperative approach to the influence of lexical neighborhood characteristics on pronunciation.

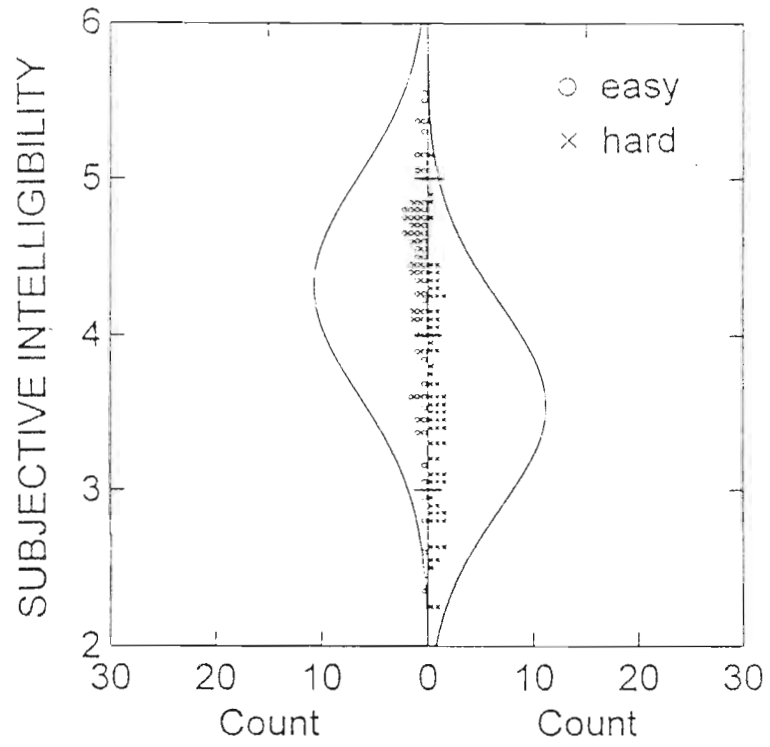


Figure 1. Distribution of 75 "easy" and 75 "hard" words along subjective intelligibility scale.

Reported Strategies

Although the subjective intelligibility ratings show that subjects can accurately predict which words should be more or less difficult to recognize in a noisy environment, they do not show that such predictions were based on explicit judgments of lexical properties such as frequency, neighborhood density, or neighborhood frequency. In order to determine the grounds for the subjective intelligibility ratings, we turned to the post-test questionnaire, in which subjects were asked to describe any strategies they had used to estimate how intelligible the target words would be in a noisy environment. Three general kinds of strategies emerged. First, *frequency*-based strategies involved judgments of how common or familiar a word was, with more frequent words being seen as more intelligible. Second, *neighborhood-density*-based strategies involved judgments of how many other words sounded similar to the target word, with words from sparser neighborhoods being seen as more intelligible. Third, *phoneme*-based strategies involved judgments of how distinctive or salient the phonemes of a word were. Without exception, subjects adopting this kind of strategy focused on consonants, and assumed that words with stops or affricates in the initial or final position would be more intelligible.

The number of subjects (out of 20) reporting each of these strategies is shown in Table 1. (Note that several subjects reported using more than one strategy.) The most common strategy involved judgments of neighborhood density, and over half of the subjects adopted this strategy. Frequency- and phoneme-based strategies were less widely used, and roughly equal numbers of subjects adopted each of these strategies. Finally, four subjects reported not having used any particular strategy to estimate the intelligibility of the target words.

Table 1

Strategies Used for Subjective Intelligibility Ratings

Strategy	Number of Subjects
<i>frequency</i>	5
<i>neighborhood density</i>	11
<i>phoneme</i>	6
<i>none</i>	4

The fact that a significant number of subjects attended to either the frequency or the neighborhood density of words during the ratings task suggests that speakers not only have access to such lexical properties, but also recognize how these properties may influence word recognition. However, this does not necessarily imply that speakers make explicit judgments of frequency or neighborhood density during speech production and modify their pronunciations accordingly. Rather, the influence of lexical neighborhood characteristics on pronunciation may often be due to implicit knowledge of the relationship between such characteristics and intelligibility. Supporting this possibility, all four of the subjects who reported not having used any particular strategy during the ratings task nevertheless gave the “easy” words higher subjective intelligibility ratings than the “hard” words.

Of course, the fact that some of the subjects reported using a phoneme-based strategy – namely, listening for stops or affricates in the initial or final position – suggests that the results of the ratings task could have been due to articulatory differences between the “easy” and “hard” words. Further inspection of the materials revealed that 30 of the “easy” words contained stops or affricates in the initial position, and 24 contained them in the final position. In contrast, 32 of the “hard” words contained stops or

affricates in the initial position, and 39 contained them in the final position. Thus, if anything, the phoneme-based strategy should have biased the subjective intelligibility ratings in favor of the “hard” words. Indeed, the single subject who gave the “hard” words higher subjective intelligibility ratings than the “easy” words reported using this strategy.

Conclusions

In this study, we found that speakers can accurately predict that words from sparse neighborhoods and having high relative usage frequencies should be more easily identified than words from dense neighborhoods and having low relative usage frequencies. This finding is consistent with the cooperative approach to the influence of lexical neighborhood characteristics on pronunciation, according to which speakers continually assess the ability of listeners to recover the message, and modify their pronunciations accordingly. However, the present results are by no means conclusive. For example, the subjective intelligibility ratings task could have tapped into metalinguistic knowledge that, although potentially useful from a cooperative perspective, plays no actual role during speech production.

To further evaluate the cooperative approach, several follow-up studies are needed in which subjects are first asked to pronounce “easy” and “hard” words, and later provide subjective intelligibility ratings for these words. Of particular interest would be whether individual differences in the degree of hyper-articulation of “hard” words are positively correlated with individual differences in the degree to which “hard” words are given lower subjective intelligibility ratings than “easy” words. That is, do those subjects showing the greatest sensitivity to the distinction between “easy” and “hard” words in the production task also show the greatest sensitivity to this distinction in the ratings task? A positive answer would lend further support to the cooperative framework.

References

- Anderson, A. H., Bard, E. G., Sotillo, C., Newlands, A., & Doherty-Sneddon, G. (1997). Limited visual control of the intelligibility of speech in face-to-face dialogue. *Perception and Psychophysics*, *39*, 580-592.
- Balota, D. A., Boland, J. E., & Shields, L. W. (1989). Priming in pronunciation: Beyond pattern recognition and onset latency. *Journal of Memory and Language*, *28*, 14-36.
- Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language*, *24*, 89-106.
- Bolinger, D. (1963). Length, vowel, juncture. *Linguistics*, *1*, 5-29.
- Carroll, J. B. (1971). Measurement properties of subjective magnitude estimates of word frequency. *Journal of Verbal Learning and Verbal Behavior*, *10*, 722-729.
- Chafe, W. (1974). Language and consciousness. *Language*, *50*, 111-133.
- Charles-Luce, J. (1987). The effects of semantic context on voicing neutralization. In *Research on Speech Perception Progress Report No. 13* (pp. 167-195). Bloomington, IN: Speech Research Laboratory, Indiana University.
- Fowler, C. A., & Housum, J. (1987). Talkers' signaling of “new” and “old” words in speech and listeners' perception and use of the distinction. *Memory and Language*, *26*, 489-504.

- Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, **113**, 256-281.
- Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, **28**, 501-518.
- Goldinger, S. D., & Summers, W. V. (1989). Lexical neighborhoods in speech production: A first report. In *Research on Speech Perception Progress Report No. 15* (pp. 331-342). Bloomington, IN: Speech Research Laboratory, Indiana University.
- Hawkins, S., & Warren, P. (1994). Phonetic influences on the intelligibility of conversational speech. *Journal of Phonetics*, **22**, 493-511.
- Hunnicut, S. (1985). Intelligibility vs. redundancy – conditions of dependancy. *Language and Speech*, **28**, 47-56.
- Lane, H. L., & Tranel, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech and Hearing Research*, **14**, 677-709.
- Lane, H. L., Tranel, B., & Sisson, C. (1970). Regulation of voice communication by sensory dynamics. *Journal of the Acoustical Society of America*, **47**, 618-624.
- Lieberman, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and Speech*, **6**, 172-188.
- Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H and H theory. In W. Hardcastle & A. Marchal (Eds.), *Speech production and speech modeling*. Dordrecht: Kluwer.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, **19**, 1-36.
- Nusbaum, H. C., Pisoni, D. B., & Davis, C. K. (1984). Sizing up the Hoosier mental lexicon: Measuring the familiarity of 20,000 words. *Research on Speech Perception Progress Report No. 10*. (pp. 357-376). Bloomington, IN: Speech Research Laboratory, Indiana University.
- Pisoni, D. B., Nusbaum, H. C., Luce, P. A., & Slowiaczek, L. M. (1985). Speech perception, word recognition, and the structure of the lexicon. *Speech Communication*, **4**, 75-95.
- Shapiro, B. J. (1969). The subjective estimation of relative word frequency. *Journal of Verbal Learning and Verbal Behavior*, **8**, 248-251.
- Torretta, G. M. (1995). The “easy-hard” word multi-talker speech database: An initial report. *Research on Spoken Language Processing Progress Report No. 20* pp. 321-334). Bloomington, IN: Speech Research Laboratory, Indiana University.