Listen carefully: The risk of error in spoken medication orders

Bruce L. Lambert [a,*], Laura Walsh Dickey [b], William M. Fisher [c], Robert D. Gibbons [a], Swu-Jane Lin [a], Paul A. Luce [d], Conor T. McLennan [e], John W. Senders [f], Clement T. Yu [a]

[a] University of Illinois at Chicago, Chicago, IL, United States
[b] University of Pittsburgh, United States
[c] Independent Consultant, United States
[d] University at Buffalo, United States
[e] Cleveland State University, United States
[f] University of Toronto, Canada

ABSTRACT

Clinicians and patients often confuse drug names that sound alike. We conducted auditory perception experiments in the United States to assess the impact of similarity, familiarity, background noise and other factors on clinicians' (physicians, family pharmacists, nurses) and laypersons' ability to identify spoken drug names. We found that accuracy increased significantly as the signal-to-noise (S/N) ratio increased, as subjective familiarity with the name increased and as the national prescribing frequency of the name increased. For clinicians only, similarity to other drug names reduced identification accuracy, especially when the neighboring names were frequently prescribed. When one name was substituted for another, the substituted name was almost always a more frequently prescribed drug. Objectively measurable properties of drug names can be used to predict confusability. The magnitude of the noise and familiarity effects suggests that they may be important targets for intervention. We conclude that the ability of clinicians and lay people to identify spoken drug names is influenced by signal-to-noise ratio, subjective familiarity, prescribing frequency, and the similarity neighborhoods of drug names.

© 2010 Elsevier Ltd. All rights reserved.

Introduction

In clinical medicine, the risks of misinterpretation of telephone orders are widely recognized (Koczmara, Jelincic, & Perri, 2006; Pennsylvania Patient Safety Authority, 2006; The Joint Commission, 2008). The use of the telephone to communicate medication orders leads to error because of both ambient noise and the limited bandwidth of most telephones (Aronson, 2004; Hoffman & Proulx, 2003; Lambert, 2008; Rodman, 2003; Wiener, Liu, Nelson, & Hoffman, 2004). Telephones typically carry signals between 300 Hz and 3 kHz, a much narrower bandwidth than that of FM radio (30 Hz–15 kHz) or CD audio (20 Hz–20 kHz), whereas much of the important acoustic information that allows people to distinguish between similar consonant sounds lies above 3 kHz and is missing entirely from the telephone signal (Rodman, 2003). There are 3.8 billion prescriptions dispensed in outpatient pharmacies annually in the United States (IMS Health, 2008). Telephone

orders account for 3–4% of retail prescription volume. This translates to 114 million telephone prescriptions annually, or 312,000 per day. One study of 813 telephone orders to two chain pharmacies found that the wrong medication name was transcribed in 1.4% of the orders (Camp, Hailemeskel, & Rogers, 2003). The 1.4% rate may not be a generalizable estimate, but given the number of opportunities, even a very low error rate would translate into a large number of errors.

Spoken orders were once common in inpatient settings also, although less so after accrediting agencies pressed for their elimination. One 346-bed hospital counted 4197 medication-related verbal orders in a seven day period (Wakefield et al., 2008). Hospital pharmacists reported 35 min of every 8 h shift were spent resolving problems with spoken orders (Allinson, Szeinbach, & Schneider, 2005). Respondents identified "people talking in the background" and "background noise" as the greatest barriers to the correct processing of spoken orders. Other factors included lack of familiarity with the patient's clinical condition or the medication, bad connections and excessively rapid speech (Allinson et al., 2005).

The use of cell phones and voicemail and the noisy environments in which orders are sent and received increase the risk of spoken prescription orders being misperceived. There are many

---

examples of auditory perception errors, some with fatal consequences (e.g., Liquibid vs. Lithobid, Cardene vs. codeine, Lopid vs. Slobid, erythromycin vs. azithromycin, Klonopin vs. clonidine, Viscerol vs. vistaril, Orgaran vs. argatroban) (Allinson et al., 2005; Dr. orders Liquibid–Lithobid dispensed – death results. Case on point: Clifford v. Geritom Med., Inc., 681 N.W.2d 680-MN (2004), 2004; Koczmara et al., 2006; Pennsylvania Patient Safety Authority, 2006; Vivian, 2004).

Identifying the factors that influence accuracy in the perception of spoken drug names may facilitate interventions designed to make telephone orders safer. The U.S. Food and Drug Administration and the pharmaceutical industry have struggled to develop methods for evaluating the confusability of new drug names (U. S. Food and Drug Administration, 2008). Objective measures of similarity and prescribing frequency can reliably predict the probability that two names will be confused in visual perception and short term memory (Lambert, 1997; Lambert, Chang, & Gupta, 2003; Lambert, Chang, & Lin, 2001b; Lambert, Donderi, & Senders, 2002; Lambert, Lin, Gandhi, & Chang, 1999; Lambert, Yu, & Thirumalai, 2004), and processes have been described for designing safer drug names (Lambert, Lin, & Tan, 2005). One important part of that process is to use established experimental paradigms from psycholinguistics to evaluate the confusability of proposed drug names in relevant tasks (e.g., auditory perception, visual perception, and short term memory). An earlier study of noise and pharmacy dispensing errors found, counter-intuitively, that noise improved performance but recommended that more controlled experiments be done to clarify the relationship between noise and error rates (Flynn et al., 1996). In this study, we sought to demonstrate how this type of experimentation could shed light on the factors that influence auditory perception of drug names. Thus, one of the key challenges we addressed was to translate the basic science of auditory word perception into the applied domain of drug name confusion. The work was designed to determine how and to what extent characteristics of drug names, order takers, and practice environments affect a listener's ability to identify spoken drug names.

Auditory word perception

To explain auditory perceptual confusions, we used Luce's Neighborhood Activation Model (NAM) (Grossberg, 1986; Luce, Goldinger, Auer, & Vitevitch, 2000; Luce & Pisoni, 1998; Vitevitch & Luce, 1999). According to this model, stimulus input activates a set of similar sounding acoustic-phonetic patterns in memory. The activation levels of the acoustic-phonetic patterns are a function of their degree of match with the input. In turn, these patterns activate a set of word decision units tuned to the acoustic-phonetic patterns. The word decision units compute probabilities for each pattern based on the intelligibility and frequency of occurrence of the word to which the pattern corresponds and the activation levels and frequencies of occurrence of all other similar sounding words in the system. The word decision unit that computes the highest probability wins, and its word is what is heard. In short, word decision units compute probability values based on the acoustic-phonetic similarity of the word to the input, the frequency of the word, and the activation levels and frequencies of all other similar words activated in memory.

The NAM predicts that multiple activation has consequences: Spoken words with many similar-sounding, higher frequency (or more commonly occurring) neighbors will be processed more slowly and less accurately than words with few neighbors. These predictions have been confirmed in many studies: Words in densely populated, high frequency similarity neighborhoods are indeed processed less quickly and less accurately than words in low density, lower frequency neighborhoods, and words with higher frequency of occurrence are processed more rapidly and accurately than lower frequency words (Jusczyk & Luce, 2002; Lambert et al., 2003).

The NAM employs an explicit mathematical function that attempts to predict auditory perceptual errors based on the intelligibility of stimulus word, the frequency of occurrence of the stimulus word, and the similarity and frequency of neighboring words. This function is known as frequency-weighted neighborhood probability (FWNP, or neighborhood probability). Detailed mathematical descriptions of the function used to compute neighborhood probabilities for each name are given elsewhere (Jusczyk & Luce, 2002; Lambert, Lin, Toh et al., 2005). Other things being equal, neighborhood probabilities will increase as the number, similarity, and prescribing frequency of neighbors decrease. The Neighborhood Activation Model provided the framework for the development of several hypotheses about auditory perception: (1) Accuracy will increase as neighborhood probability increases; (2) Accuracy will increase as the signal-to-noise (S/N) ratio increases; (3) Accuracy will increase as objective prescribing frequency of the target name increases; and (4) Accuracy will increase as subjective familiarity with the target name increases.

Although frequency and neighborhood effects are well established in the study of ordinary words, we sought to extend this understanding in three ways. First, we planned to study proper names from a large, finite set of words, namely the closed-set lexicon of drug names, given that most previous work has been done with open-set word lists (Clopper, Pisoni, & Tierney, 2006; Sommers, Kirk, & Pisoni, 1997). Second, we wished to study multisyllabic words rather than the monosyllabic (often consonant-vowel-consonant) words that have typically been used in studies of neighborhood effects. That required us to develop measures of similarity and new measures of neighborhood probabilities for multisyllabic words (Lambert, Lin, Toh et al., 2005). Third, we wished to see whether the neighborhood effects would be present in both experts (clinicians) and novices (laypeople). We expected neighborhood effects in experts because, presumably, they would possess lexical representations (i.e., mental word lists) of large numbers of drug names, and these representations would compete in the manner described above. We thought neighborhood effects might be attenuated or absent in laypeople, who may lack representations for most drug names and hence would not experience the competition that causes neighborhood effects.

Even if we were not making any original contribution to the basic science of auditory perception, we believe that the model provides a powerful conceptual and experimental framework for understanding drug name confusion, one that could advance worldwide efforts to predict and prevent such errors. Thus, the present paper is offered as an example of translational research, where concepts well-known to one community are applied to problems in a different domain (Woolf, 2008).

We attempted to control for a large set of factors that might be associated with name confusion error rates. Among these was the type of name (brand or generic). One might expect brand names to be more confusing since they are typically shorter (Lambert, Chang, & Lin, 2001a), and shorter names tend to have more neighbors (Andrews, 1997; Luce & Pisoni, 1998; Storker, 2004). Conversely, generic names use a common system of stems (i.e., suffixes) which tends to increase their average similarity to one another, thereby increasing their confusability (Lambert et al., 2001a). Either way, the distinction between brand and generic names is an important one in practice, so we designed our experiments to take it into account.

Materials and methods

Design

We used a within-participants design to examine the effect of signal-to-noise (S/N) ratio, prescribing frequency, subjective familiarity, and similarity neighborhoods on the ability to correctly identify spoken drug names. All participants heard an equal number of drug names at all three S/N ratios and all of the drug names appeared at all three S/N ratios. Three versions of the experiment ensured that drug names and S/N ratios were counterbalanced across participants and that participants only heard one S/N ratio version of each drug.

Participants

Participants were paid for participating. The experiment was approved by the institutional review boards at the University of Illinois at Chicago and Cleveland State University. Sixty-seven pharmacists were recruited at the 2005 meeting of the American Pharmacists Association, 76 family physicians at the 2005 meeting of the American Association of Family Physicians, and 74 nurses at the 2005 meeting of the Academy of Medical Surgical Nurses. Due to equipment malfunction or non-native language accents, 5 pharmacists, 2 physicians, and 4 nurses were excluded, leaving $N = 62$ pharmacists, $N = 74$ physicians and $N = 70$ nurses for the main analysis. Forty-three lay people were paid in exchange for their participation, and were recruited from the Cleveland State University community by flyers advertising paid language experiments for people 18 years and older and for whom American English was the first language they learned to speak, were right handed, and had no history of speech or hearing disorders. In fact, 60 laypeople were recruited, but funds were available to transcribe only the first 43 sets of responses for the main analysis. All clinicians were licensed and actively practicing at the time of the data collection. Non-native speakers of American English, left-handed individuals,[1] and anyone with serious speech or hearing problems were excluded (Gonzalez & McLennan, 2007; Hagmann et al., 2006).

Stimulus materials

We selected 99 brand and 99 generic drug names. Names and prescribing frequency information were drawn from four sources: (1) the National Ambulatory Medical Care Survey (National Center for Health Statistics, 2001a). (2) the National Hospital Ambulatory Medical Care Survey (National Center for Health Statistics, 2001b), (3) the IMS-Health National Prescription Drug Audit (IMS, 2001), and (4) the Solucient Hospital Drug Utilization Database (Solucient Inc., 2003).

A frequency-weighted neighborhood probability was computed for each name, according to a procedure described elsewhere (Lambert, Lin, Toh et al., 2005). Names were stratified by neighborhood probability, with ten brand and ten generic names from each decile of probability. One brand name and one generic name were removed to make the total evenly divisible into three S/N conditions. The names were recorded by a trained phonetician/phonologist. These *reference pronunciations* were

based on phonemic transcriptions from experienced clinicians, collected for a different project. Each recorded name was edited into an AIFF audio file. To mimic the bandwidth limitations of telephone audio, frequencies below 300 Hz and above 3 kHz were then digitally filtered (Rodman, 2003).

Stimulus degradation

Drug names were played back against a background of standard 20-speaker babble (obtained from Auditec of St. Louis). The noise was played at a mean amplitude of 65 dB and was not bandwidth limited (Flynn et al., 1996). The stimuli were played at either 63 dB, 68 dB, or 73 dB, resulting in three signal-to-noise conditions of −2 dB, +3 dB, and +8 dB respectively. During pilot testing on practicing pharmacists, these signal-to-noise ratios produced error rates of roughly 25%, 50%, and 75% at the low, medium and high levels. A subgroup of 10 lay participants completed the experiment "in the clear" with no added noise.

Experimental procedure

For clinicians, letters were mailed to attendees in advance, and individuals were approached at the convention center and asked: (a) if they were licensed, practicing clinicians in the U.S. and (b) if they were interested in participating in a study of drug name confusion. Lay people were recruited as described above in the Participants section. Other than differences in recruitment, the experimental procedures were identical for lay and clinical participants. Participants consented, completed a demographic questionnaire, and took a pure-tone hearing threshold test (pure tone thresholds of 50 dB or lower were accepted). Participants were seated at a Macintosh PowerBook computer and fitted with headphones with an attached microphone (Beyerdynamic BT190). The participant then read the instructions. Playback of the 20-speaker babble was initiated and continued for the duration of the experiment.

The PsyScope experiment program (Cohen, MacWhinney, Flatt, & Provost, 1993) was used to run the main experiment. The task began with 21 practice trials and continued with 198 trials in random order as the main experiment. One name was subsequently dropped from the analysis due to an error in recording. On each trial, participants were asked to repeat back the name they had heard. Spoken responses were recorded through the laptop's built in microphone. After completing the main experiment, participants moved to another computer and read aloud the 198 experimental names as they were visually presented on a computer screen. Participants rated their familiarity with each name on a 7-point scale (extremely familiar − extremely unfamiliar).

Scoring spoken responses

Spoken responses were transcribed into the ARPAbet (Jurafsky & Martin, 2000). phonetic alphabet by an experienced linguist (e.g., Zyvox was transcribed as /Z ′1AY V ′2AA K S/, where the numerals "1" and "2" indicate stress level). Responses were scored as correct or incorrect by comparing the transcribed responses to the reference transcriptions for each of the 197 stimulus names. Because variation in pronunciation made verbatim matching to the reference transcription too strict a criterion, we developed additional procedures to capture legitimate pronunciation variants. The first was a computer program that applied generally accepted rules for pronunciation variation to the reference pronunciations. For example, in fluent speech, unstressed vowel sounds are reduced to the schwa sound (schwa is a short neutral vowel sound, the most common vowel sound in English, e.g. the first phoneme in the word/again/). Responses were scored as correct if they could be produced by applying the variation rules to the reference

---

[1] It is customary to exclude left-handed participants from language research because right-handers typically represent and process language in their left hemisphere, and there is more variability in how left-handers represent and process language. Hemispheric differences could have consequences for the speed and/or accuracy with which language is processed. So, to reduce unnecessary noise in the data, left-handers are often excluded.

pronunciations. Even after applying these rules, there still appeared to be legitimate variants that were being scored as incorrect. So, the linguist examined all incorrectly scored responses, identified those deemed to be legitimate variants, and provided linguistic justification for each case. In the end, a response was scored as correct if it matched the reference pronunciation exactly, if it could be automatically generated as a rule-generated variant, or if it was recognized by our expert in phonetics as a legitimate variant.

Analysis plan

The goal of our analysis was to quantify the main effects of neighborhood probability, noise, frequency and familiarity on accuracy in auditory perception. Descriptive analyses were done using SAS version 9.1, and to test the hypotheses, mixed effects logistic regression models were built with SuperMix (Hedeker & Gibbons, 2008).

The dependent variable was accuracy. Words correctly identified were scored as 1, and words incorrectly identified were scored as 0. The independent variables were: (1) the neighborhood probability, a continuous variable on the interval 0 to 1, reflecting the *predicted* probability of identification; (2) signal to noise ratio, an ordinal variable with three levels (−2 dB, +3 dB, +8 dB); (3) familiarity, an ordinal variable reflecting a participant's subjective familiarity with the name (ranging from 1 to 7); (4) prescribing frequency, a continuous variable representing the log (base 10) of the maximum frequency from our multiple sources of prescribing frequency data; (5) phoneme frequency, a continuous variable representing the frequency of a given consonant or vowel in a given position in the word; (6) biphone frequency, a continuous variable representing the frequency of a two-phoneme sequence in a given position in the word (Storker, 2004). The control variables were (1) participant demographics, including age, gender, race, practice context, profession and years of experience; (2) pure tone threshold, eight continuous variables reflecting the sensitivity of a participant's hearing in each ear at 500 Hz, 1 kHz, 2 kHz and 3 kHz; (3) length, an ordinal variable reflecting the number of phonemes in the word; and (4) trial, an ordinal variable representing the sequential position of a given response within the set of 198 responses.

To test our hypotheses, we built a logistic regression model for the combined group of clinicians (pharmacists, physicians and nurses) and a separate model for lay people, treating the intercepts as a random effect. We also carried out subgroup analyses for each clinical profession, the results of which are presented in selected tables and figures where space permits. We identified the correct scale for each independent and control variable by plotting the log odds of error as a function of each variable. If these plots were linear, terms were entered as linear. If the plot revealed an obvious nonlinearity, we selected a scale to fit the nonlinear form of the function (Hosmer & Lemeshow, 1989; Selvin, 1996). We used Kleinbaum's method of backward elimination to decide which variables to include in the final model (Kleinbaum, 1994). This method begins with a full model and proceeds to eliminate as many terms as possible, using likelihood ratio tests to decide which terms contribute significantly to the model's fit. All statistical tests used $p < 0.05$. Finally, we examined the fit between observed and predicted accuracy rates for each of the 197 stimulus names. Observed rates were taken directly from the data. For each participant, we used the parameter estimates from the fitted model and covariate values from the data to generate a predicted proportion of correct identifications for each name averaged across all participants.

Results

Each of the 239 participants responded to 197 stimuli, producing 47,083 total responses. Mean accuracy on the perception task was 32.27% (s.d. = 7.94, range = 12.18%–49.75%, also see Table 5). Roughly 77% of correct responses were identified by verbatim matching, 9.5% by computer scoring of alternative pronunciations and 13.5% by expert scoring of alternatives. This error rate is roughly two orders of magnitude greater than what one would expect in real-world practice (Flynn, Barker, & Carnahan, 2003). The task was intentionally designed to be difficult and to produce high error rates because (a) we were interested in studying the errors themselves, and we wanted a large sample of errors to analyze, and (b) due to statistical power considerations, the naturalistic error rate (perhaps 0.13%) (Flynn et al., 2003) would have made it impossible to detect any differences in error rate across experimental conditions. The task was not designed to estimate real-world error rates (that is best achieved by direct observation). Rather, our goal was to understand the factors that affected the error rate in a task that was analogous to, but more difficult than, the real-world task.

Table 1 describes participant demographics. Table 2 gives the means and standard deviations of continuous predictors for correct and incorrect responses as well as results of bivariate tests of association between each predictor and the log odds of a correct response. Table 2 provides the initial indication of the beneficial effects of familiarity and prescribing frequency on accuracy for both clinicians and laypeople. For clinicians, frequency-weighted neighborhood probability (FWNP) and word length were also associated with increased accuracy. For laypeople, neighborhood probability had no impact, and longer names were actually perceived less accurately than shorter names. The number of phonemes helped clinicians because each additional phoneme provided disambiguating information that aided them in discriminating between similar alternatives in memory. For laypeople, most of whom were unfamiliar with the stimulus names, the task was really a bottom-up identification task, and additional phonemes made that task harder. A similar analysis could be made in relation to neighborhood probability. Probability predicted clinician performance because probability measures the extent to which a name is free of competition from similar neighbors. But since laypeople lacked mental representations for most of these words, neighborhood probability was not a relevant measure for them. Individual difference variables like hearing acuity, age or experience had little or no impact on accuracy (especially for clinicians), whereas properties of the environment or the stimulus names did.

Table 3 gives the proportion of correct and incorrect responses in relation to the nominal predictors, as well as results of tests of the bivariate association between predictors and accuracy. For clinicians, accuracy more than tripled from the noisiest to the quietest condition. In laypeople, the effect was even stronger. For clinicians, gender, type of name (brand vs. generic), race and practice context were significantly associated with accuracy in bivariate analyses, but many of these effects were confounded with type of clinician and were not significant in multivariate analyses.

Fig. 1 illustrates the powerful effect of familiarity on accuracy. Performance on the most familiar names was roughly three times better than on the least familiar names. This is a consequence of the well-known word frequency effect, which predicts that familiar words will be perceived more accurately than unfamiliar words due either to higher resting activation levels or to decision biases that favor familiar items in memory (Howes, 1957). Fig. 2 shows the equally dramatic effects of noise on accuracy, again with performance improving by a factor of three for clinicians and by a factor of more than 5 for lay people as S/N ratio increased from −2 dB to +8 dB. Fig. 2 also shows that without noise, lay people performed as well as pharmacists in the +8 dB condition. There was a clear gradient in accuracy, with pharmacists being most accurate, followed by physicians, nurses and lay people. As might be expected,

Table 1
Demographic characteristics.

Variable	Pharmacists ($n = 62$)		Physicians ($n = 74$)		Nurses ($n = 70$)		Lay people ($n = 43$)	
Age M (SD)	39.1 (10.7)		44.3 (10.9)		43.2 (7.7)		29.8 (12.3)	
	N	%	N	%	N	%	N	%
Gender								
Male	26	40.9	50	67.6	1	1.4	11	25.6
Female	36	58.1	24	32.4	69	98.6	32	74.4
Race								
Asian	3	4.8	16	21.9	1	1.4	0	0
Pacific islander	1	1.6	0	0	0	0	3	7.1
Black	5	8.1	2	2.7	0	0	11	26.2
White	50	80.7	54	74.0	66	94.3	24	57.1
Multiracial	1	1.6	0	0	1	1.4	4	9.5
Other	2	3.2	1	1.4	2	2.9	1	0.2
Work context								
Hospital	7	11.3	6	8.1	69	98.6	–	–
Clinic	13	21.0	61	82.4	1	1.4	–	–
Retail	33	53.2	0	0	0	0	–	–
Community	9	14.5	0	0	0	0	–	–
Other	0	0.0	7	9.5	0	0	–	–

performance appeared to be determined by each group's familiarity with drug names – greater familiarity led to higher accuracy.

Table 4 gives the parameter estimates for the final random effect logistic regression models. Clinicians and laypeople produced different patterns of results, presumably because clinicians had mental representations of many of the names, but laypeople did not. For clinicians, the task involved (bottom-up) identification and (top-down) discrimination among similar competing alternatives. For laypeople, it was primarily a bottom-up identification task. Signal-to-noise ratio was associated with significantly improved accuracy in participants' identification of drug names. For clinicians, neighborhood probability was associated with increased accuracy. Names with fewer and less frequently prescribed neighbors were heard more accurately than names with greater numbers of more frequently prescribed neighbors. For both lay people and clinicians, familiar drug names were perceived more accurately than unfamiliar names, although the relationship was nonlinear for

clinicians. For clinicians, but not for laypeople, names with high prescribing frequency were perceived more accurately than names with low prescribing frequency. For clinicians, accuracy increased as the number of phonemes per name increased while the opposite was true for laypeople (for the reason described above). For both clinicians and laypeople, biphone frequency was positively associated with accuracy. For clinicians the effect of biphone frequency was modified somewhat by phoneme frequency. Thus, the presence of common sound patterns made names easier to recognize. For clinicians, S/N ratio interacted significantly with familiarity, as did phoneme and biphone frequency.

Clinicians got slightly better at the task as they completed more trials. Brand name drugs were perceived less accurately by clinicians than generic names, even after controlling for name length and number of neighbors. In subgroup analyses (details not shown) this effect was restricted to physicians. Physicians and nurses had lower accuracy scores than pharmacists. For laypeople, age and

Table 2
Mean (standard deviation) of continuous independent variables for correct and incorrect responses.

Variable	Clinicians ($N = 206$)			Laypeople ($N = 33$)		
	Incorrect ($n = 26776$)	Correct ($n = 13806$)	P	Incorrect ($n = 5111$)	Correct ($n = 1390$)	P
R500 Hz	25.92 (8.32)	26.07 (8.60)	0.44	12.94 (4.74)	12.65 (4.87)	0.26
L500 Hz	24.12 (10.91)	23.92 (10.53)	0.37	11.91 (4.56)	11.50 (4.63)	0.09
R1000 Hz	17.72 (7.10)	17.43 (7.06)	0.05	8.02 (5.93)	8.07 (6.35)	0.87
L1000 Hz	16.03 (7.45)	15.86 (7.42)	0.26	6.92 (5.73)	6.45 (5.83)	0.14
R2000 Hz	10.85 (8.75)	10.71 (8.17)	0.38	4.35 (6.41)	3.83 (6.40)	0.13
L2000 Hz	10.89 (8.97)	10.64 (8.49)	0.15	4.66 (6.76)	4.83 (6.63)	0.57
R3000 Hz	10.96 (10.47)	11.14 (10.12)	0.45	3.37 (8.19)	2.49 (7.94)	0.04
L3000 Hz	11.72 (11.10)	11.73 (10.89)	0.95	5.20 (8.19)	4.97 (7.86)	0.65
Phoneme Freq.	−0.11 (1.19)	−0.01 (1.10)	<0.01	−0.14 (1.21)	0.07 (1.01)	<0.01
Biphone Freq.	−0.14 (1.03)	0.04 (1.02)	<0.01	−0.14 (1.03)	0.08 (1.01)	<0.01
Age	42.43 (10.13)	42.18 (10.06)	0.25	30.62 (12.42)	30.41 (11.86)	0.81
Experience	14.79 (10.19)	14.63 (10.15)	0.42	–	–	–
Familiarity	2.73 (2.36)	4.57 (2.65)	<0.01	1.61 (1.44)	2.63 (2.28)	<0.01
FWNP	0.53 (0.34)	0.60 (0.36)	<0.01	0.55 (0.35)	0.56 (0.35)	0.34
Rx Frequency	3.14 (1.65)	3.94 (1.76)	<0.01	3.37 (1.75)	3.58 (1.65)	<0.01
No. Phonemes	8.06 (2.29)	8.25 (2.80)	<0.01	8.25 (2.53)	7.67 (2.22)	<0.01
Trial	121.95 (57.13)	123.62 (57.22)	0.01	99.40 (547.09)	100.22 (57.39)	0.63

Note. R500 Hz is the lowest number of decibels at which a 500 Hz tone could still be heard in the right ear. L500 Hz is the left ear, etc. FWNP is frequency-weighted neighborhood probability. Rx Frequency is the log (base 10) of the national prescribing frequency of the stimulus names. Mixed effects logistic regression models were built with Supermix with only an intercept and one independent variable in the model. P values come from Wald tests on the parameter estimate for the variable in question. Lay results are for $n = 33$ lay particpants in the condition with background noise. There were 197 observations for each participant, so there were 12,214 pharmacist responses, 14,578 physician responses, and 13,790 nurse responses. See text for details.

Table 3
Percent correct and incorrect responses by nominal covariates.

Variable	Clinicians ($N = 206$)			Lay People ($N = 33$)		
	% Cor.	% Inc.	P	% Cor.	% Inc.	P
S/N Ratio			<0.0001			<0.0001
−2 dB	16.03	83.97		6.83	93.17	
+3 dB	36.10	63.90		21.32	78.68	
+8 dB	49.93	50.07		35.99	64.01	
Gender			<0.0001			0.31
Male	36.09	63.91		23.13	76.87	
Female	32.78	67.22		20.91	79.09	
Name type			<0.0001			0.83
Brand	31.30	68.70		21.27	78.73	
Generic	36.76	63.24		21.49	78.51	
Race			<0.0001			0.09
White	34.44	65.56		22.73	77.27	
Non-white	32.05	67.95		19.76	80.24	
Context			<0.0001			–
Hospital	30.28	69.72		–	–	–
Clinic	35.42	64.58		–	–	–
Retail	38.70	61.30		–	–	–
Other	36.88	63.12		–	–	–

Note. Probabilities (P) come from Wald tests on parameter estimates generated by SuperMix mixed effects regression models with only one predictor variable entered at a time. Lay results are for $n = 33$ lay particpants in the condition with background noise. A dash indicates that varible was not included in the analysis.

hearing acuity were associated with accuracy. For clinicians, only left ear pure tone threshold at 2000 Hz was associated with accuracy.

To assess goodness of fit of the model, we compared the observed and predicted percent correct for each of the 197 drug names. The root mean square error of prediction ranged from 15.3% to 18.5% (see Table 5). Several names were rarely identified correctly (e.g., sutilains, *Kira*, sparfloxacin, and tromethamine), while others were rarely missed (e.g., hydrochlorothiazide, *Zithromax*, codeine). The models accounted for between 32% and 49% of the variance in accuracy at the level of the individual name.

Substitution errors

The most common type of error was an incorrect pronunciation of the stimulus name (90%), but one in ten errors was a *substitution error*, where the response corresponded to another real drug name. Substitutions produce potentially harmful wrong-drug errors. Models of spoken word recognition predict that rare names will be misheard as common names but not vice versa. We tested this hypothesis. Overwhelmingly, substitution errors went in the direction of the more frequently prescribed name. There were 4692 substitution errors overall. In 2963 (63.2%) of these errors, the substituted name had a higher prescribing frequency than the stimulus name. In 1411 cases (30.1%), the substituted name was not in our prescribing frequency databases. Considering only the cases with known prescribing frequency, errors went in the direction of the more frequently prescribed drug 90.31% of the time. For clinicians, the mean difference in log prescribing frequency between the substituted name and the target name was 2.40 (s.d. = 1.82), meaning the substituted name was, on average, prescribed 250 times more frequently than the target. For errors that went in the direction of the more frequently prescribed name, the mean difference in log frequency was 2.75. For errors that went in the direction of the lower frequency name, the mean difference in log prescribing frequency was only 1.22. This suggests that confusions go in the direction of the less common name only when the prescribing frequencies are relatively similar. As the difference in prescribing frequency increases so does the tendency for the errors to go in the direction of the higher frequency name (see Fig. 3).

Discussion

Our goal was to identify factors that affect accuracy in auditory perception of drug names. Results of our experiments supported our three specific hypotheses and our general model of auditory perception: (1) Accuracy (for clinicians) was influenced by the similarity neighborhood of each drug name, i.e., by the similarity and prescribing frequencies of neighboring names. Names with less similar and less commonly prescribed neighbors were more accurately perceived than names with more similar and more
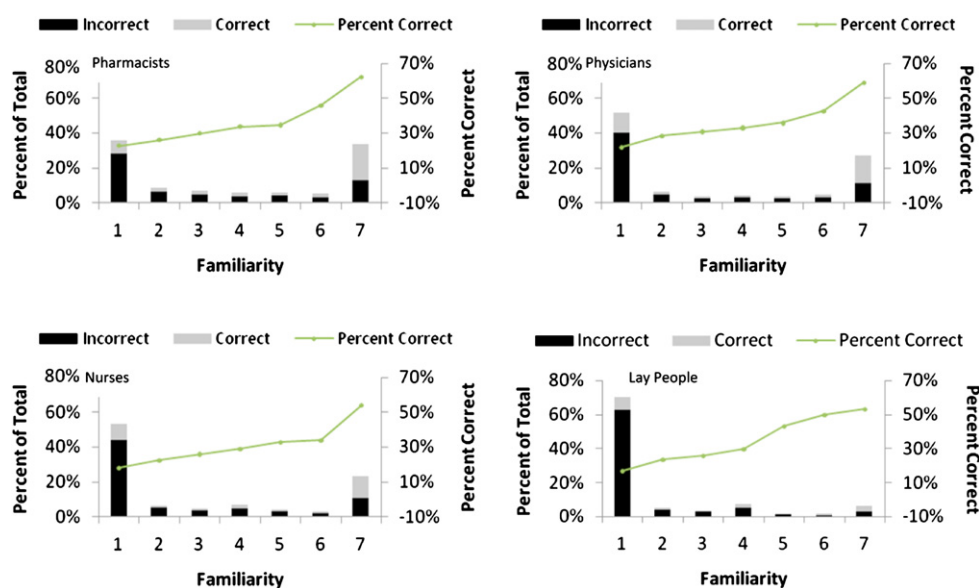

Fig. 1. Auditory perception accuracy as a function of familiarity. For each level of familiarity, the dark bar represents the percent of incorrect responses as a percent of the total number of responses for that participant group. The light bar represents the percent of correct responses at that familiarity level. The line represents the percent correct at a given level of familiarity. For lay people, data are shown only from participants in the condition with background noise.
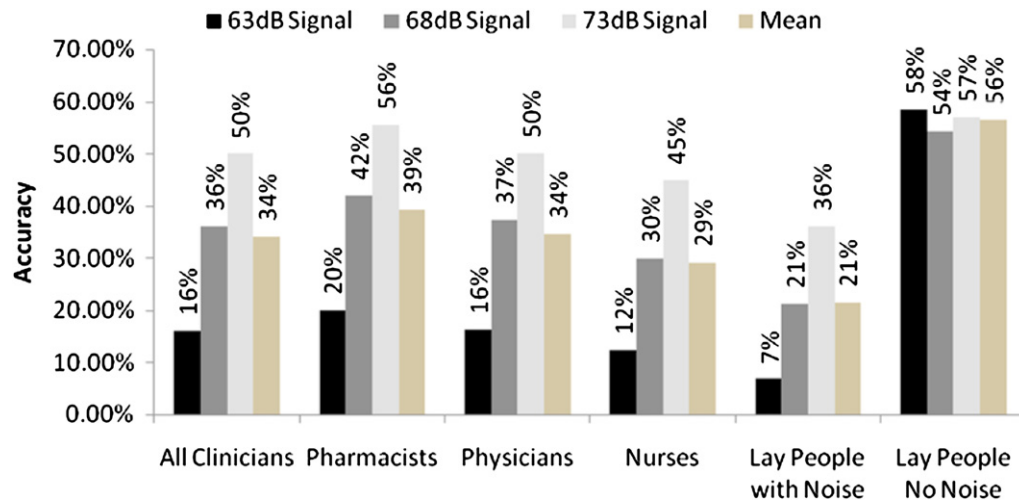
Fig. 2. Effect of signal strength on accuracy for pharmacists, physicians, nurses and lay people. Noise was played at mean 65 dB amplitude for all participant groups, except where noted for lay people. So the three signal-to-noise conditions corresponded to −2 dB, +3 dB and +8 dB. Lay people were the only subgroup tested without noise.

frequently prescribed neighbors. Names with fewer and less frequently prescribed neighbors were subject to less competition during word recognition and were therefore perceived more accurately. (2) Accuracy increased as S/N ratio increased. And (3)

familiar, more frequently prescribed names were perceived more accurately than unfamiliar, less frequently prescribed names.

Some of these associations were more complex than we had predicted. The relationship between neighborhood probability and

Table 4
Parameter estimates for random effects logistic regression model of accuracy in auditory perception for clinicians and lay people.

Variable	Clinicians ($n = 206$)		Lay People ($n = 33$)	
	OR (95% CI)	P	OR (95% CI)	P
Intercept	0.21 (0.15–0.30)	<0.0001	0.03 (0.1–0.8)	<0.0001
S/N Ratio (dB)	1.23 (1.22–1.24)	<0.0001	1.23 (1.21–1.25)	<0.0001
FWNP	0.83 (0.41–1.68)	0.60	–	–
FWNP^2	0.31 (0.06–1.60)	0.16	–	–
FWNP^3	4.34 (1.51–12.53)	0.01		
Rx Frequency	0.99 (0.94–1.04)	0.69	–	–
Rx Frequency^2	1.02 (1.01–1.02)	<0.0001	–	–
Familiarity	2.09 (1.65–2.66)	<0.0001	1.40 (1.35–1.46)	<0.0001
Familiarity^2	0.83 (0.78–0.90)	<0.0001	–	
Familairity^3	1.02 (1.01–1.02)	<0.0001	–	
Num. Phonemes	0.82 (0.78–0.86)	<0.0001	0.92 (0.89–0.94)	<0.0001
Num. Phonemes^2	1.01 (1.01–1.01)	<0.0001	–	
Phoneme Freq.	0.73 (0.68–0.78)	<0.0001	–	
Biphone Freq.	1.69 (1.58–1.82)	<0.0001	1.29 (1.20–1.37)	<0.0001
Familiarity × S/N Ratio	0.995 (0.993–0.998)	0.0001	–	–
Familiarity × Phoneme Freq.	1.04 (1.03–1.06)	<0.0001	–	–
Familiarity × Biphone Freq.	0.95 (0.93–0.96)	<0.0001	–	–
Nonwhite	0.79 (0.70–0.89)	0.0001	–	–
Trial	1.0009 (1.0005–1.0013)	<0.0001	–	–
Brand	0.93 (0.88–0.98)	0.01	–	–
Physician	0.91 (0.82–1.01)	0.08	–	–
Nurse	0.66 (0.59–0.74)	<0.0001	–	–
Age			1.11 (1.04–1.17)	0.0007
Age^2			0.999 (0.998–0.999)	0.002
R500 Hz			0.96 (0.94–0.99)	0.003
R1000 Hz			1.03 (1.01–1.05)	0.01
R2000 Hz			0.97 (0.95–0.99)	0.01
L2000 Hz	0.99 (0.98–0.99)	<0.0001	1.02 (1.00–1.04)	0.03
R3000 Hz			0.98 (0.96–0.99)	0.02
L3000 Hz			1.02 (1.00–1.04)	0.02

Note. FWNP = frequency-weighted neighborhood probability. Rx Frequency is the log base 10 of the maximum frequency observed across 4 different prescription databases. S/N ratio is the signal-to-noise ratio. For race, the reference group was and white (Caucasian). R500 Hz is the lowest number of decibels at which a 500 Hz tone could still be heard in the right ear. L500 Hz is the left ear, etc. A dash means the variable was not significant in the model. Superscript numbers are exponents, e.g., FWNP^2 = FWNP squared. See text for details.

Table 5
Summary of drug name identification results for pharmacists, physicians, nurses and laypeople.

Variable	All clinicians ($n = 206$)	Pharmacists ($n = 62$)	Physicians ($n = 74$)	Nurses ($n = 70$)	Laypeople ($n = 33$)
Accuracy (%)					
Mean (s.d.)	34.23 (6.82)	39.14 (5.83)	34.46 (5.52)	29.02 (5.15)	21.38 (5.36)
Range	13.71–49.75	16.75–49.75	19.8–47.2	13.71–38.58	12.18–30.46
Goodness of fit					
Root mean squared error	17.23%	18.5%	17.5%	17.8%	15.3%
Mean absolute error	14.22%	15.5%	14.7%	14.6%	11.7%
R2 predicted vs. observed	0.47	0.49	0.48	0.42	0.32
Substitution errors					
Number (% of all responses)	4449 (10.96)	1780 (14.6)	1404 (9.66)	1265 (9.17)	243 (3.73)
% of incorrect responses	16.62	23.95	14.69	12.92	4.75
No. with known freq.	3128	1266	975	887	153
No. (%) in direction of higher frequency name	2849 (91.1)	1163 (91.9)	883 (90.6)	803 (90.5)	114 (74.5)
Mean (s.d.) log freq. difference between target and substituted name	2.40 (1.80)	2.57 (1.80)	2.27 (1.76)	2.30 (1.83)	1.54 (2.06)

Note. For laypeople, data are shown only for participants in the condition with background noise.

accuracy was quadratic for clinicians, and the relationship between familiarity and accuracy was cubic for clinicians. In both cases this is likely because some words, such as the low familiarity names, were simply unknown to the participants. These unknown words were processed as if they were non-words, and the effects of many of our predictors on words and non-words are known to differ. For analogous reasons, we saw interactions between familiarity and other predictors, such as S/N ratio, phoneme frequency, and biphone frequency. These relationships need to be explored more thoroughly in subsequent work.

Our models accounted for a substantial amount of the observed variance in item-level accuracy. We believe such models would be useful to policy makers and drug companies as they evaluate the confusability of new and existing drug names. We found confusions to be asymmetrical, with substitutions overwhelmingly going in the direction of the more frequent name. Thus, when policy makers consider the potential for harm related to a name confusion, they must consider the direction of the error (i.e., which drug will be mistaken for which). More familiar and more frequently prescribed drugs will tend to have lower error rates, but in thinking about
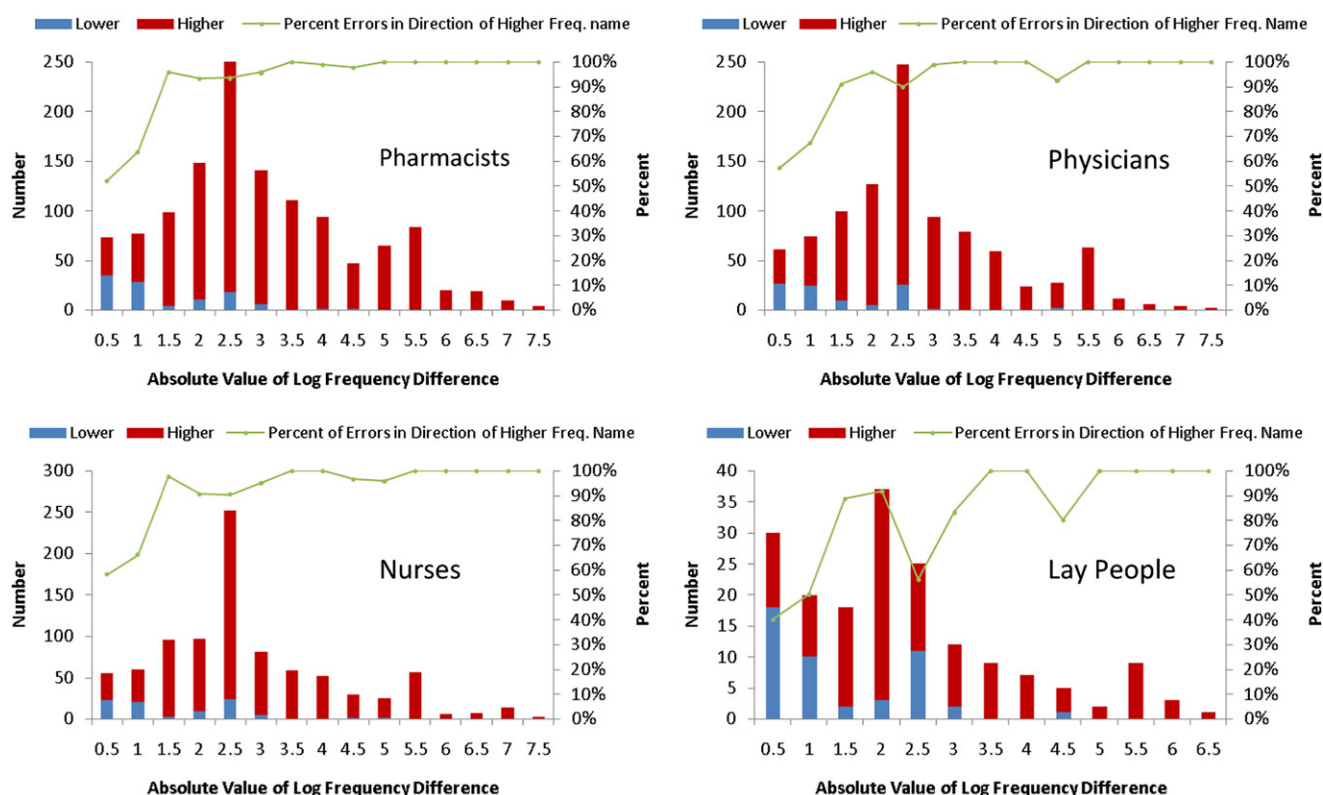

Fig. 3. Direction of substitution errors as a function of difference in log prescribing frequency. The bar chart is a histogram of frequency differences between stimulus names and substituted names. The bottom part of each vertical column represents the number of times the substituted name was less frequently prescribed than the stimulus name (left axis). The top portion represents the number of times the substituted name was more frequently prescribed than the stimulus name. The line represents the percent of substitution errors wherein the substituted name was more frequently prescribed than the stimulus name (right axis). The graphs show that when one name is mistaken for another, the substituted name is almost always more frequently prescribed than the name which was misheard. The probability of error is not a symmetrical function of similarity. Relatively low frequency names are liable to be misheard as their higher frequency neighbors, but not vice versa.

harm, this must be weighed against their higher number of opportunities for error (Lambert et al., 2003).

Limitations

We measured subjective familiarity after the participants completed the auditory perception task. Therefore, it is possible that our measure of subjective familiarity may have been influenced by experience in the experiment itself. Although possible, we expect that this effect, if it existed at all, was small, especially in comparison to the long-term priming effects of lifetime experience (or lack thereof) with these stimulus names. Previous work on subjective familiarity suggests that it is an accurate measure of lifetime exposure to words and that it is better than objective frequency for predicting word recognition performance (Balota, Pilotti, & Cortese, 2001; Carroll, 1971; Gernsbacher, 1984).

We studied a convenience sample of right-handed, native English-speaking clinicians and lay people. We studied only 197 drug names out of perhaps 11,000 or more drug names in use in the U.S. Generalizations to other clinicians and other drug names should be made cautiously. The level and type of noise we used may not reflect real working conditions, especially because the noise was of a constant type and amplitude. Real world contexts experience unpredictable noise of varying types and intensities. The lack of foreign language accents, distractions, interruptions and of clinical context was also unrealistic. This work should be replicated under more realistic conditions.

Conclusion

Objectively measurable properties of drug names can be used to predict their confusability. The ability to accurately identify spoken drug names is influenced by signal-to-noise ratio, subjective familiarity, prescribing frequency, and the similarity neighborhoods of drug names. To minimize errors, order takers should be able to increase the source volume, and should have noise-cancelling headphones or quiet areas where they can take orders. Although not directly supported by these experiments, it is generally agreed that spoken communication of drug names can be made safer by using strategies like read-back, spelling out the name, providing the indication for the drug or using both brand and generic names (especially for physicians who were more likely to misperceive brand than generic names). The finding that laypeople, with no background noise, achieved recognition scores about equal to those achieved by expert clinicians operating under moderate signal-to-noise ratios, underscores the importance of reducing local and remote background noise and giving order takers the ability to increase voice signal loudness. Since unfamiliar drug names are more likely to be misperceived than familiar names, it may be possible to reduce drug name confusions by increasing clinicians' familiarity with a wider array of drug names. It may also be possible to improve performance by having people learn the distinguishing parts of similar drug names through repeated discrimination training. However, the effectiveness of both suggested interventions could have unintended consequences and have yet to be evaluated. The experiments and models described above should prove useful to regulatory agencies and drug manufacturers who must evaluate the confusability of drug names prior to allowing them on the market.

Acknowledgements

References

Allinson, T. T., Szeinbach, S. L., & Schneider, P. J. (2005). Perceived accuracy of drug orders transmitted verbally by telephone. *American Journal of Health-System Pharmacy, 62*, 78–83.

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: resolving neighborhood conflicts. *Psychonomic Bulletin and Review, 4*(4), 439–461.

Aronson, J. K. (2004). Medication errors resulting from the confusion of similar names. *Expert Opinion on Drug Safety, 3*, 167–172.

Balota, D., Pilotti, M., & Cortese, M. (2001). Subjective frequency estimates for 2938 monosyllabic words. *Memory and Cognition, 29*(4), 639–647.

Camp, S. C., Hailemeskel, B., & Rogers, T. L. (2003). Telephone prescription errors in two community pharmacies. *American Journal of Health-System Pharmacy, 60*, 613–614.

Carroll, J. B. (1971). Measurement properties of subjective magnitude estimates of word frequency. *Journal of Verbal Learning and Verbal Behavior, 10*, 722–729.

Clopper, C., Pisoni, D., & Tierney, A. (2006). Effects of open-set and closed-set task demands on spoken word recognition. *Journal of the American Academy of Audiology, 17*(5), 331–349.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: a new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25*(2), 257–271.

Dr. orders liquibid–lithobid dispensed − death results. Case on point: Clifford v. Geritom Med., Inc., 681 N.W.2d 680-MN (2004). (2004). *Nursing Law's Regan Report, 45*(2), 4.

Flynn, E. A., Barker, K. N., & Carnahan, B. J. (2003). National observational study of prescription dispensing accuracy and safety in 50 pharmacies. *Journal of the American Pharmacists Association (Washington), 43*(2), 191–200.

Flynn, E. A., Barker, K. N., Gibson, J. T., Pearson, R. E., Smith, L. A., & Berger, B. A. (1996). Relationships between ambient sounds and accuracy of pharmacists' prescription-filling performance. *Human Factors, 38*, 614–622.

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.

Gonzalez, J., & McLennan, C. (2007). Hemispheric differences in indexical specificity effects in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 410–424.

Grossberg, S. (1986). The adaptive self-organization of serial order in behavior: speech, language and motor control. In E. C. Schwab, & H. C. Nusbaum (Eds.), *Pattern recognition by humans and machines: Speech perception, Vol. 1* (pp. 187–294). New York: Academic Press.

Hagmann, P., Cammoun, L., Martuzzi, R., Maeder, P., Clarke, S., Thiran, J.-P., et al. (2006). Hand preference and sex shape the architecture of language networks. *Human Brain Mapping, 27*, 828–835.

Hedeker, D., & Gibbons, R. D. (2008). SuperMix mixed effects models. from. http://www.ssicentral.com/supermix/index.html.

Hoffman, J. M., & Proulx, S. M. (2003). Medication errors caused by confusion of drug names. *Drug Safety, 26*(7), 445–452.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: John Wiley & Sons.

Howes, D. H. (1957). On the relation between the intelligibility and frequency of occurrence of English words. *Journal of the Acoustical Society of America, 29*, 296–305.

IMS Health. (2008). 2007 channel distribution by U.S. dispensed prescriptions. Retrieved November 17, 2008, from. http://www.imshealth.com/deployedfiles/imshealth/Global/Content/Document/Top-Line%20Industry%20Data/2007%20Channel%20Distribution%20by%20RXs.pdf.

IMS, I. (2001). *National prescription audit plus*. Chicago, IL: IMS. Inc.

Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice Hall.

Jusczyk, P. W., & Luce, P. A. (2002). Speech perception. In H. Pashler, & S. Yantis (Eds.) (3rd. ed.).*Stevens' handbook of experimental psychology: Sensation and perception, Vol. 1* (pp. 493–536) New York: John Wiley and Sons.

Kleinbaum, D. G. (1994). *Logistic regression: A self-learning text*. New York: Springer-Verlag.

Koczmara, C., Jelincic, V., & Perri, D. (2006). Communication of medication orders by telephone-"writing it right". *Dynamics, 17*(1), 20–24.

Lambert, B. L. (1997). Predicting look- and sound-alike medication errors. *American Journal of Health-System Pharmacy, 54*, 1161–1171.

Lambert, B. L. (2008). Recent developments in the prevention and detection of drug name confusion. In R. W. Hicks, S. C. Becker, & D. D. Cousins (Eds.), *MEDMARX data report. A report on the relationship of drug names and medication errors in response to the Institute of Medicine's call for action* (pp. 10–15). Rockville, MD: Center for the Advancement of Patient Safety, US Pharmacopeia.

Lambert, B. L., Chang, K.-Y., & Gupta, P. (2003). Effects of frequency and similarity neighborhoods on pharmacists' visual perception of drug names. *Social Science & Medicine, 57*, 1939–1955.

Lambert, B. L., Chang, K. Y., & Lin, S. J. (2001a). Descriptive analysis of the drug name lexicon. *Drug Information Journal, 35*, 163–172.

Lambert, B. L., Chang, K. Y., & Lin, S. J. (2001b). Effect of orthographic and phonological similarity on false recognition of drug names. *Social Science & Medicine, 52*, 1843–1857.

Lambert, B. L., Donderi, D., & Senders, J. (2002). Similarity of drug names: objective and subjective measures. *Psychology and Marketing, 19*(7–8), 641–661.

Lambert, B. L., Lin, S.-J., Gandhi, S. K., & Chang, K.-Y. (1999). Similarity as a risk factor in drug name confusion errors: the look-alike (orthographic) and sound-alike (phonological) model. *Medical Care, 37*(12), 1214–1225.

Lambert, B. L., Lin, S.-J., & Tan, H. K. (2005). Designing safe drug names. *Drug Safety, 28*, 495–512.

Lambert, B. L., Lin, S.-J., Toh, S., Luce, P. A., McLennan, C. T., La Vigne, R., et al. (2005). Frequency and neighbourhood effects on auditory perception of drug names in noise. Retrieved from. *Acoustical Society of America Journal, 118*(3), 1955. http://tigger.uic.edu/~lambertb.con_proceed/n05_173.pdf.

Lambert, B. L., Yu, C., & Thirumalai, M. (2004). A system for multi-attribute drug product comparison. *Journal of Medical Systems, 28*(1), 29–54.

Luce, P. A., Goldinger, S. D., Auer, E. T., Jr., & Vitevitch, M. S. (2000). Phonetic priming, neighborhood activation, and PARSYN. *Perception and Psychophysics, 62*, 615–625.

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

National Center for Health Statistics. (2001a). Ambulatory health care data: NAMCS description. Retrieved May 31, 2001, from. http://www.cdc.gov/nchs/about/major/ahcd/namcsdes.htm.

National Center for Health Statistics. (2001b). Ambulatory health care data: NHAMCS description. Retrieved May 31, 2001, from. http://www.cdc.gov/nchs/about/major/ahcd/nhamcsds.htm.

Pennsylvania Patient Safety Authority. (2006). Improving the safety of telephone or verbal orders. Retrieved from. *Patient Safety Advisory, 3*(2), 1. http://www.psa.state.pa.us/psa/lib/psa/advisories/v3n2june2006/junevol_3_no_2_article_b_safety_of_telphone_orders.pdf 3–7.

Rodman, J. (2003). The effect of bandwidth on speech intelligibility. Retrieved June 17, 2005, from. http://www.polycom.com/common/pw_cmp_updateDocKeywords/0,1687,1809,00.pdf.

Selvin, S. (1996). *Statistical analysis of epidemiological data*. New York: Oxford University Press.

Solucient Inc.. (2003). Hospital drug utilization database. Retrieved March 5, 2004, from. http://www.solucient.com/solutions/Solucients_Databases.shtml.

Sommers, M. S., Kirk, K. I., & Pisoni, D. B. (1997). Some considerations in evaluating spoken word recognition by normal-hearing, noise-masked normal hearing and cochlear implant users I: the effects of response format. *Ear and Hearing, 18*, 89–99.

Storker, H. L. (2004). Methods for minimizing the confounding effects of word length in the analysis of phonotactic probability and neighborhood density. *Journal of Speech, Language, and Hearing Research, 47*, 1454–1468.

The Joint Commission. (2008). Accreditation program: ambulatory health care, national patient safety goals. Retrieved from. http://www.jointcommission.org/NR/rdonlyres/979098FA-74FD-4F25-AF41-EDD48FBD300E/0/AHC_NPSG.pdf.

U.S. Food and Drug Administration. (2008). PDUFA pilot project proprietary name review. *Federal Register, 74*(189), 50806–50809.

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language, 40*, 374–408.

Vivian, J. V. (2004). Third-party actions. *US Pharmacist, 11*, 74–78.

Wakefield, D. S., Ward, M. M., Groath, D., Schwichtenberg, T., Magdits, L., Brokel, J., et al. (2008). Complexity of medication-related verbal orders. *American Journal of Medical Quality, 23*(1), 7–17.

Wiener, S. W., Liu, S., Nelson, L. S., & Hoffman, R. S. (2004). Nutropin or Neupogen? A medication error resulting in leukocytosis. *Annals of Pharmacotherapy, 38*(4). 721–721.

Woolf, S. (2008). The meaning of translational research and why it matters. *Journal of the American Medical Association, 299*(2), 211.