

**Frequency Distributions for Names and Unconstrained Words
Associated with the Letters of the English Alphabet**

TR 29.3437

July 21, 2001

James R. Lewis

IBM Voice Systems

West Palm Beach, Florida

Abstract

This report provides frequency distributions for words and names associated with the letters of the English alphabet. These distributions can be useful when developing voice spelling user interfaces. The data suggest that the unconstrained word distribution is preferable to the names distribution because, contrary to expectation, the overall consistency of unconstrained responses was not poorer than the responses in the names distribution and the occurrence of cases in which participants were not able to provide a response for the names distribution was significantly greater than for the unconstrained distribution.

ITIRC Keywords

Voice spelling
Word choice distributions
Name choice distributions
Unconstrained word distributions
Name constrained distributions

Contents

Introduction	1
Method	3
Participants	3
Materials	3
Procedure	3
Results	5
Frequency Distributions for Each Letter of the Alphabet	5
Correlation Analyses	55
Comparison of Unconstrained and Name Distributions	55
Discussion	57
References	59

Introduction

Voice spelling is an important component of many speech systems (embedded in other appliances, accessed by telephone, or allows hands-free use of a desktop dictation system). If a user needs to use a word that is not in the system's built-in lexicon and either cannot or chooses not to use a keyboard, then the user must spell the word¹ by voice.

The standard approach to voice spelling is to let users say the letter names or, because some letter names have high acoustic confusability, to also permit the use of the military alphabet (alpha, beta, Charlie, etc.). The military alphabet has an advantage over letter names in that the words chosen for the military alphabet have very low acoustic confusability. The military alphabet has a disadvantage over letter names in that it is very difficult to memorize the military alphabet.

The purpose of the current research was to find out what words come first to people's minds when asked to produce a word for each letter of the alphabet. In one experiment participants provided words without constraint. In a separate experiment, participants provided first names for each letter. The reason for constraining participants in this way was to see if participants under this condition produced more consistent responses than those in the unconstrained condition.

¹ Note that this is true for alphanumeric strings that, strictly speaking, are not words, such as PIN codes, part numbers, and membership numbers.

Method

Participants

The source for participants in each experiment was a set of 400 IBM employees (800 employees in all, 400 for each experiment), selected at random from an internal e-mail directory of all of the IBM employees in the United States and sent an e-mail invitation to participate. Of the employees invited to participate, 103 (26%) responded to the invitation for the first experiment (no constraints) and 120 (30%) responded to the invitation for the second experiment (names for letters).

Materials

I used the IBM User-Centered Design's version of WebSurveyor² to construct a web-based form for the evaluation. The form simply provided a space by each letter of the alphabet (arranged vertically on the page in alphabetical order), with instructions appropriate for the specific experiment (to either provide any word for each letter, or to provide only first names).

Procedure

After receiving the e-mailed invitation, participants clicked a link in the message that brought up the WebSurveyor page containing the survey that the participant was to take. The participants read an introduction explaining the purpose of the survey, then completed and submitted the form. The WebSurveyor program kept track of the raw data. After collection, I used a spreadsheet to organize and summarize the results.

² WebSurveyor is a trademark of WebSurveyor Corp.

Results

The following tables contain the results of the surveys, organized by letter of the alphabet. Each table contains the submitted words, listed in decreasing frequency of submission. The table also shows the cumulative frequency and indicates (1) the point at which the cumulative frequency for a given frequency of submission came as close as possible to 80%³ (indicated by an asterisk), (2) the number of words included in the list up to that point (beside the cell with the asterisk), (3) the standard deviation of the frequencies, and (4) the total number of different words submitted [with (3) and (4) respectively appearing in the last two cells of each table].

Frequency Distributions for Each Letter of the Alphabet

Table 1. Frequency distribution for A (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
apple	67.0	67.0		
alpha	11.7	78.6		
able	4.9	83.5	*	3
Adam	3.9	87.4		
alphabet	1.9	89.3		
achieve	1.0	90.3		
adobe	1.0	91.3		
albatross	1.0	92.2		
Albert	1.0	93.2		
Alice	1.0	94.2		
analysis	1.0	95.1		
angry	1.0	96.1		
animal	1.0	97.1		
announce	1.0	98.1		
around	1.0	99.0	St Dev	Count
arrive	1.0	100.0	16.4	16

³ The choice of 80% as a cutoff point was arbitrary, but seemed reasonable for the purpose of ensuring fairly comprehensive coverage and for comparing the cutoff points of the different distributions. The tables allow other interested researchers to work with any other cutoff points.

Table 2. Frequency distribution for A (first names)

Name	Percent	Cumulative	Cutoff	Count
Adam	10.8	10.8		
Alan	9.2	20.0		
Albert	5.8	25.8		
Alex	5.8	31.7		
Alice	5.8	37.5		
Alpha	5.8	43.3		
Ann	5.8	49.2		
Apple	5.8	55.0		
Anna	5.0	60.0		
Abel	4.2	64.2		
Andrew	4.2	68.3		
Al	3.3	71.7		
Amy	2.5	74.2		
Andy	2.5	76.7		
Anthony	2.5	79.2	*	15
Aaron	1.7	80.8		
Alfred	1.7	82.5		
Amanda	1.7	84.2		
April	1.7	85.8		
Adams	0.8	86.7		
Ailing	0.8	87.5		
Aires	0.8	88.3		
Alana	0.8	89.2		
Alvin	0.8	90.0		
Ana	0.8	90.8		
Andrea	0.8	91.7		
Angela	0.8	92.5		
Animal	0.8	93.3		
Anita	0.8	94.2		
Annie	0.8	95.0		
Art	0.8	95.8		
Arthur	0.8	96.7		
Ashley	0.8	97.5		
Ashwini	0.8	98.3		
Athena	0.8	99.2	St Dev	Count
Audra	0.8	100.0	2.6	36

Table 3. Frequency distribution for B (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
boy	40.4	40.4		
baker	9.6	50.0		
bravo	9.6	59.6		
ball	4.8	64.4		
bear	3.8	68.3		
butterfly	2.9	71.2		
baby	1.9	73.1		
banana	1.9	75.0		
beta	1.9	76.9		
bingo	1.9	78.8		
bob	1.9	80.8		
bottom	1.9	82.7		
box	1.9	84.6		
boxer	1.9	86.5	*	14
bag	1.0	87.5		
baseball	1.0	88.5		
beautiful	1.0	89.4		
bee	1.0	90.4		
beer	1.0	91.3		
bewildered	1.0	92.3		
big	1.0	93.3		
bill	1.0	94.2		
biography	1.0	95.2		
bird	1.0	96.2		
brain	1.0	97.1		
brown	1.0	98.1		
bump	1.0	99.0	St Dev	Count
business	1.0	100.0	7.6	28

Table 4. Frequency distribution for B (first names)

Name	Percent	Cumulative	Cutoff	Count
Bob	33.3	33.3		
Brian	8.3	41.7		
Barbara	6.6	48.3		
Baker	5.8	54.1		
Betty	5.8	59.9		
Boy	5.8	65.8		
Beta	3.3	69.1		
Bill	3.3	72.4		
Bobbie	3.3	75.8		
Bravo	3.3	79.1	*	10
Bonnie	2.5	81.6		
Brad	2.5	84.1		
Becky	1.7	85.8		
Ben	1.7	87.4		
Baohua	0.8	88.3		
Barb	0.8	89.1		
Bart	0.8	89.9		
Beatrice	0.8	90.8		
Berta	0.8	91.6		
Beth	0.8	92.4		
Blake	0.8	93.3		
Boat	0.8	94.1		
Boris	0.8	94.9		
Bradley	0.8	95.8		
Brandon	0.8	96.6		
Brava	0.8	97.4		
Bruce	0.8	98.3		
Bunty	0.8	99.1	St Dev	Count
Butch	0.8	99.9	6.1	29

Table 5. Frequency distribution for C (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
cat	38.5	38.5		
Charley	37.5	76.0		
Charles	2.9	78.8	*	3
car	1.9	80.8		
card	1.9	82.7		
call	1.0	83.7		
camel	1.0	84.6		
candle	1.0	85.6		
candy	1.0	86.5		
caravan	1.0	87.5		
Carl	1.0	88.5		
carrot	1.0	89.4		
catalogue	1.0	90.4		
change	1.0	91.3		
child	1.0	92.3		
cloud	1.0	93.3		
color	1.0	94.2		
company	1.0	95.2		
confused	1.0	96.2		
contract	1.0	97.1		
convert	1.0	98.1		
corporation	1.0	99.0	St Dev	Count
cup	1.0	100.0	10.6	23

Table 6. Frequency distribution for C (first names)

Name	Percent	Cumulative	Cutoff	Count
Charlie	32.3	32.3		
Cathy	12.8	45.1		
Charles	7.5	52.6		
Carol	6.8	59.4		
Carl	6.0	65.4		
Chris	5.3	70.7		
Cat	3.8	74.4		
Chuck	3.0	77.4	*	8
Candi	1.5	78.9		
Catherine	1.5	80.5		
Cindy	1.5	82.0		
Connie	1.5	83.5		
Cailian	0.8	84.2		
Calvin	0.8	85.0		
Candace	0.8	85.7		
Carlos	0.8	86.5		
Caroline	0.8	87.2		
Cecilia	0.8	88.0		
Charlotte	0.8	88.7		
Chia	0.8	89.5		
Christian	0.8	90.2		
Chrisy	0.8	91.0		
Chunnu	0.8	91.7		
Clark	0.8	92.5		
Claude	0.8	93.2		
Cody	0.8	94.0		
Cole	0.8	94.7		
Colin	0.8	95.5		
Color	0.8	96.2		
Cooper	0.8	97.0		
Cynthia	0.8	97.7	St Dev	Count
C: None	2.3	100.0	6.0	32

Table 7. Frequency distribution for D (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
dog	66.0	66.0		
David	13.6	79.6	*	2
delta	8.7	88.3		
deal	1.0	89.3		
default	1.0	90.3		
Denmark	1.0	91.3		
dependancy	1.0	92.2		
determination	1.0	93.2		
diaper	1.0	94.2		
different	1.0	95.1		
distraught	1.0	96.1		
diversity	1.0	97.1		
Don	1.0	98.1		
donut	1.0	99.0	St Dev	Count
dot	1.0	100.0	16.8	15

Table 8. Frequency distribution for D (first names)

Name	Percent	Cumulative	Cutoff	Count
David	45.8	45.8		
Dave	9.2	55.0		
Dog	7.5	62.5		
Doug	5.8	68.3		
Daniel	4.2	72.5		
Delta	4.2	76.7		
Don	4.2	80.8	*	7
Deborah	1.7	82.5		
Dennis	1.7	84.2		
Diane	1.7	85.8		
Donald	1.7	87.5		
Dad	0.8	88.3		
Dahlgren	0.8	89.2		
Dan	0.8	90.0		
Danny	0.8	90.8		
Daviv	0.8	91.7		
Dean	0.8	92.5		
Debbie	0.8	93.3		
Diabetes	0.8	94.2		
Dick	0.8	95.0		
Dominique	0.8	95.8		
Dong Fan	0.8	96.7		
Donna	0.8	97.5		
Dyurodhan	0.8	98.3	St Dev	Count
D: None	1.7	100.0	9.0	25

Table 9. Frequency distribution for E (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
elephant	21.7	21.7		
Edward	14.2	35.8		
echo	12.3	48.1		
easy	10.4	58.5		
egg	6.6	65.1		
eagle	3.8	68.9		
effort	2.8	71.7		
extra	2.8	74.5		
eat	1.9	76.4		
every	1.9	78.3		
excellent	1.9	80.2	*	11
ear	0.9	81.1		
Earl	0.9	82.1		
Easter	0.9	83.0		
eel	0.9	84.0		
Einstein	0.9	84.9		
elbow	0.9	85.8		
element	0.9	86.8		
elite	0.9	87.7		
emergency	0.9	88.7		
empty	0.9	89.6		
energy	0.9	90.6		
enough	0.9	91.5		
enterprise	0.9	92.5		
equal	0.9	93.4		
Erving	0.9	94.3		
essay	0.9	95.3		
estimate	0.9	96.2		
evening	0.9	97.2		
event	0.9	98.1		
ever	0.9	99.1	St Dev	Count
exray	0.9	100.0	4.8	32

Table 10. Frequency distribution for E (first names)

Name	Percent	Cumulative	Cutoff	Count
Edward	29.4	29.4		
Ed	7.6	37.0		
Eric	5.0	42.0		
Echo	4.2	46.2		
Ellen	4.2	50.4		
Easy	3.4	53.8		
Elephant	3.4	57.1		
Edgar	2.5	59.7		
Elmer	2.5	62.2		
Emily	2.5	64.7		
Ernie	2.5	67.2		
Eugene	2.5	69.7		
Earnie	1.7	71.4		
Eddie	1.7	73.1		
Edith	1.7	74.8		
Elizabeth	1.7	76.5		
Evan	1.7	78.2		
Evelyn	1.7	79.8	*	19
Eagle	0.8	80.7		
Earl	0.8	81.5		
Early	0.8	82.4		
Earnest	0.8	83.2		
Eat	0.8	84.0		
Edison	0.8	84.9		
Egg	0.8	85.7		
Eintein	0.8	86.6		
Eliza	0.8	87.4		
Elsa	0.8	88.2		
Elvira	0.8	89.1		
Emanuel	0.8	89.9		
Erica	0.8	90.8		
Ermei	0.8	91.6		
Ernest	0.8	92.4		
Esha	0.8	93.3		
Ethan	0.8	94.1		
Eve	0.8	95.0		
Evette	0.8	95.8		
Evie	0.8	96.6	St Dev	Count
E: None	3.4	100.0	4.7	39

Table 11. Frequency distribution for F (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Frank	34.0	34.0		
fox	21.4	55.3		
foxtrot	4.9	60.2		
frog	3.9	64.1		
fun	3.9	68.0		
fish	2.9	70.9		
fast	1.9	72.8		
final	1.9	74.8		
fine	1.9	76.7		
Fred	1.9	78.6		
fruit	1.9	80.6	*	11
facilitate	1.0	81.6		
facility	1.0	82.5		
fall	1.0	83.5		
fancy	1.0	84.5		
fastball	1.0	85.4		
father	1.0	86.4		
fearful	1.0	87.4		
fill	1.0	88.3		
film	1.0	89.3		
finance	1.0	90.3		
finish	1.0	91.3		
fire	1.0	92.2		
flag	1.0	93.2		
flower	1.0	94.2		
foot	1.0	95.1		
for	1.0	96.1		
forest	1.0	97.1		
forget	1.0	98.1		
Francis	1.0	99.0	St Dev	Count
front	1.0	100.0	6.8	31

Table 12. Frequency distribution for F (first names)

Name	Percent	Cumulative	Cutoff	Count
Frank	63.8	63.8		
Fred	11.8	75.6		
Fox	4.7	80.3	*	3
Foxtrot	2.4	82.7		
Frances	2.4	85.0		
Fay	1.6	86.6		
Florence	1.6	88.2		
Fatima	0.8	89.0		
Felix	0.8	89.8		
Forrest	0.8	90.5		
Frederick	0.8	91.3		
Fulu	0.8	92.1		
Fun	0.8	92.9	St Dev	Count
F: None	7.1	100.0	16.6	14

Table 13. Frequency distribution for G (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
George	25.0	25.0		
girl	16.3	41.3		
good	9.6	51.0		
golf	6.7	57.7		
green	5.8	63.5		
go	4.8	68.3		
goat	3.8	72.1		
gary	2.9	75.0		
great	2.9	77.9	*	9
God	1.9	79.8		
gold	1.9	81.7		
grape	1.9	83.7		
gamma	1.0	84.6		
garage	1.0	85.6		
Gayle	1.0	86.5		
geography	1.0	87.5		
giants	1.0	88.5		
giraffe	1.0	89.4		
gnash	1.0	90.4		
gopher	1.0	91.3		
gound	1.0	92.3		
grammer	1.0	93.3		
grapple	1.0	94.2		
grass	1.0	95.2		
greeting	1.0	96.2		
grey	1.0	97.1		
ground	1.0	98.1		
grumpy	1.0	99.0	St Dev	Count
guard	1.0	100.0	5.3	29

Table 14. Frequency distribution for G (first names)

Name	Percent	Cumulative	Cutoff	Count
George	48.8	48.8		
Gary	9.9	58.7		
Greg	9.9	68.6		
Gail	4.1	72.7		
Georgia	2.5	75.2		
Gina	2.5	77.7	*	6
Gabriel	1.7	79.3		
Gerry	1.7	81.0		
Gertrude	1.7	82.6		
Golf	1.7	84.3		
Good	1.7	86.0		
Grace	1.7	87.6		
Gregory	1.7	89.3		
Gamma	0.8	90.1		
Ganga	0.8	90.9		
Gene	0.8	91.7		
Girl	0.8	92.6		
Glenda	0.8	93.4		
Gloria	0.8	94.2		
Goal	0.8	95.0		
Grape	0.8	95.9		
Gretchen	0.8	96.7		
Grey	0.8	97.5		
Guo	0.8	98.3		
Gweneth	0.8	99.2	St Dev	Count
G: None	0.8	100.0	9.5	26

Table 15. Frequency distribution for H (unconstrained)

Word	Percent	Cumulative	c/o sd	Count
help	13.6	13.6		
Henry	13.6	27.2		
Harry	11.7	38.8		
hat	7.8	46.6		
hello	7.8	54.4		
happy	6.8	61.2		
hotel	5.8	67.0		
high	4.9	71.8		
hot	3.9	75.7		
Harold	2.9	78.6		
helicopter	2.9	81.6	*	11
hammer	1.9	83.5		
horse	1.9	85.4		
house	1.9	87.4		
hair	1.0	88.3		
heart	1.0	89.3		
heather	1.0	90.3		
heavy	1.0	91.3		
hell	1.0	92.2		
hemisphere	1.0	93.2		
hill	1.0	94.2		
homely	1.0	95.1		
hope	1.0	96.1		
horrendous	1.0	97.1		
host	1.0	98.1		
hurt	1.0	99.0	St Dev	Count
hydrant	1.0	100.0	4.0	27

Table 16. Frequency distribution for H (first names)

Name	Percent	Cumulative	Cutoff	Count
Harry	32.5	32.5		
Henry	25.8	58.3		
Harold	5.0	63.3		
Helen	5.0	68.3		
Howard	4.2	72.5		
Hank	2.5	75.0		
Hariet	2.5	77.5		
Heather	2.5	80.0		
Heidi	2.5	82.5		
Hotel	2.5	85.0	*	10
Harvey	1.7	86.7		
Hilda	1.7	88.3		
Hillary	1.7	90.0		
Hugo	1.7	91.7		
Hari	0.8	92.5		
Hazel	0.8	93.3		
Heart	0.8	94.2		
Hello	0.8	95.0		
Holly	0.8	95.8		
Hong	0.8	96.7		
Hot	0.8	97.5		
Hubert	0.8	98.3	St Dev	Count
H: None	1.7	100.0	8.0	23

Table 17. Frequency distribution for I (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
igloo	13.5	13.5		
India	12.5	26.0		
Indian	7.7	33.7		
ink	5.8	39.4		
indigo	3.8	43.3		
idiot	3.8	47.1		
ice	3.8	51.0		
island	2.9	53.8		
lda	2.9	56.7		
ice cream	2.9	59.6		
Ivan	1.9	61.5		
iron	1.9	63.5		
inside	1.9	65.4		
Ingrid	1.9	67.3		
income	1.9	69.2		
in	1.9	71.2		
ignore	1.9	73.1		
idaho	1.9	75.0		
icon	1.9	76.9	*	19
IBM	1.0	77.9		
ivy	1.0	78.8		
itch	1.0	79.8		
it	1.0	80.8		
is	1.0	81.7		
iowa	1.0	82.7		
intelligent	1.0	83.7		
instruction	1.0	84.6		
inn	1.0	85.6		
Information	1.0	86.5		
index	1.0	87.5		
incorporation	1.0	88.5		
incident	1.0	89.4		
impossible	1.0	90.4		
immediate	1.0	91.3		
imagine	1.0	92.3		
Illinois	1.0	93.3		
ill	1.0	94.2		
ike	1.0	95.2		
ignorant	1.0	96.2		
ignite	1.0	97.1		
IEW	1.0	98.1		
icky	1.0	99.0	St Dev	Count
iceberg	1.0	100.0	2.8	43

Table 18. Frequency distribution for I (first names)

Name	Percent	Cumulative	Cutoff	Count
Irene	11.1	11.1		
Ian	10.3	21.4		
Isabel	9.4	30.8		
Ida	7.7	38.5		
Ivan	6.0	44.4		
Ingrid	5.1	49.6		
Isaac	5.1	54.7		
India	4.3	59.0		
Indian	2.6	61.5		
Ichabod	1.7	63.2		
Indu	1.7	65.0		
Ines	1.7	66.7		
Iris	1.7	68.4		
Irving	1.7	70.1		
Ice	0.9	70.9		
Iggy	0.9	71.8		
Ike	0.9	72.6		
Ilo	0.9	73.5		
Imogene	0.9	74.4		
Imus	0.9	75.2		
Income	0.9	76.1		
Indigo	0.9	76.9		
Inger	0.9	77.8		
Ink	0.9	78.6		
Inn	0.9	79.5		
Insect	0.9	80.3		
Insomnia	0.9	81.2		
Intro	0.9	82.1		
Iola	0.9	82.9		
Ira	0.9	83.8		
Irv	0.9	84.6		
Isabella	0.9	85.5		
Island	0.9	86.3		
Italics	0.9	87.2		
Ivy	0.9	88.0	*	35
			St Dev	Count
I: None	12.0	100.0	3.3	36

Table 19. Frequency distribution for J (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Jack	16.5	16.5		
jump	11.7	28.2		
John	9.7	37.9		
Juliet	4.9	42.7		
James	3.9	46.6		
join	2.9	49.5		
joke	2.9	52.4		
joker	2.9	55.3		
juice	2.9	58.3		
Julie	2.9	61.2		
junk	2.9	64.1		
Japan	1.9	66.0		
jar	1.9	68.0		
Jerry	1.9	69.9		
Jim	1.9	71.8		
jolly	1.9	73.8		
joy	1.9	75.7	*	17
jack-o-lantern	1.0	76.7		
Jacob	1.0	77.7		
Jake	1.0	78.6		
jam	1.0	79.6		
january	1.0	80.6		
Java	1.0	81.6		
jacks	1.0	82.5		
Jeffrey	1.0	83.5		
jelly	1.0	84.5		
jellybean	1.0	85.4		
jerk	1.0	86.4		
jewel	1.0	87.4		
jinxed	1.0	88.3		
job	1.0	89.3		
jock	1.0	90.3		
Joe	1.0	91.3		
joint	1.0	92.2		
Jordan	1.0	93.2		
Joseph	1.0	94.2		
journey	1.0	95.1		
jumping	1.0	96.1		
june	1.0	97.1		
Jupiter	1.0	98.1		
jury	1.0	99.0	St Dev	Count
just	1.0	100.0	3.1	42

Table 20. Frequency distribution for J (first names)

Name	Percent	Cumulative	Cutoff	Count
Jack	25.6	25.6		
John	23.1	48.7		
James	6.8	55.6		
Joe	6.8	62.4		
Jim	5.1	67.5		
Jennifer	4.3	71.8		
Juliet	3.4	75.2	*	7
Jacob	1.7	76.9		
Jake	1.7	78.6		
Jane	1.7	80.3		
Janet	1.7	82.1		
Jason	1.7	83.8		
Jeff	1.7	85.5		
Jenny	1.7	87.2		
Jackie	0.9	88.0		
Jackson	0.9	88.9		
Jamie	0.9	89.7		
Jasper	0.9	90.6		
Jay	0.9	91.5		
Jayson	0.9	92.3		
Jesus	0.9	93.2		
Jhumi	0.9	94.0		
Jill	0.9	94.9		
Jing	0.9	95.7		
Joanne	0.9	96.6		
Joke	0.9	97.4		
Jony	0.9	98.3		
Joseph	0.9	99.1	St Dev	Count
Jump	0.9	100.0	6.1	29

Table 21. Frequency distribution for K (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
king	19.6	19.6		
kite	16.7	36.3		
kilo	10.8	47.1		
kitchen	5.9	52.9		
Kate	3.9	56.9		
kangaroo	2.9	59.8		
Kevin	2.9	62.7		
kick	2.9	65.7		
kid	2.9	68.6		
kind	2.9	71.6		
Kansas	2.0	73.5		
Karen	2.0	75.5		
Kathy	2.0	77.5		
kill	2.0	79.4		
killer	2.0	81.4		
kiss	2.0	83.3		
kitty	2.0	85.3	*	17
kappa	1.0	86.3		
karate	1.0	87.3		
keep	1.0	88.2		
Keisha	1.0	89.2		
Kelly	1.0	90.2		
kept	1.0	91.2		
Kevlar	1.0	92.2		
kickball	1.0	93.1		
Kim	1.0	94.1		
kindhearted	1.0	95.1		
kitten	1.0	96.1		
klepto	1.0	97.1		
knight	1.0	98.0		
knights	1.0	99.0	St Dev	Count
knock	1.0	100.0	4.4	32

Table 22. Frequency distribution for K (first names)

Name	Percent	Cumulative	Cutoff	Count
Kevin	16.2	16.2		
Karen	14.3	30.5		
Ken	13.3	43.8		
Kelley	9.5	53.3		
King	5.7	59.0		
Kilo	4.8	63.8		
Kate	2.9	66.7		
Katie	2.9	69.5		
Kim	2.9	72.4		
Kyle	2.9	75.2		
Kaitlyn	1.9	77.1		
Keith	1.9	79.0		
Kent	1.9	81.0		
Kerry	1.9	82.9	*	14
Kalpana	1.0	83.8		
Kangaroo	1.0	84.8		
Kappa	1.0	85.7		
Karlien	1.0	86.7		
Kelvin	1.0	87.6		
Kendall	1.0	88.6		
Kenneth	1.0	89.5		
Kenny	1.0	90.5		
Kill	1.0	91.4		
Kitchen	1.0	92.4		
Kite	1.0	93.3		
Kristina	1.0	94.3		
Kun	1.0	95.2		
Kurt	1.0	96.2	St Dev	Count
K: None	3.8	100.0	4.3	29

Table 23. Frequency distribution for L (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Larry	22.3	22.3		
love	10.7	33.0		
Lima	6.8	39.8		
lion	4.9	44.7		
laugh	3.9	48.5		
lost	3.9	52.4		
ladder	2.9	55.3		
lady	2.9	58.3		
life	2.9	61.2		
like	2.9	64.1		
long	2.9	67.0		
lamb	1.9	68.9		
lemon	1.9	70.9		
Linda	1.9	72.8		
little	1.9	74.8		
lollipop	1.9	76.7		
low	1.9	78.6	*	17
lake	1.0	79.6		
lamp	1.0	80.6		
lane	1.0	81.6		
lap	1.0	82.5		
large	1.0	83.5		
Laura	1.0	84.5		
lawn	1.0	85.4		
left	1.0	86.4		
lego	1.0	87.4		
letter	1.0	88.3		
lettuce	1.0	89.3		
leverage	1.0	90.3		
Lil	1.0	91.3		
lincoln	1.0	92.2		
Lisa	1.0	93.2		
listen	1.0	94.2		
log	1.0	95.1		
Lolita	1.0	96.1		
loop	1.0	97.1		
loosen	1.0	98.1		
lose	1.0	99.0	St Dev	Count
lump	1.0	100.0	3.8	39

Table 24. Frequency distribution for L (first names)

Name	Percent	Cumulative	Cutoff	Count
Larry	40.0	40.0		
Linda	11.7	51.7		
Laura	4.2	55.8		
Lewis	4.2	60.0		
Lisa	4.2	64.2		
Lou	3.3	67.5		
Lima	2.5	70.0		
Lori	2.5	72.5		
Louie	2.5	75.0		
Louise	2.5	77.5		
Lynn	2.5	80.0	*	11
Leon	1.7	81.7		
Lucy	1.7	83.3		
Lumber	0.8	84.2		
Lake	0.8	85.0		
Laugh	0.8	85.8		
Lauren	0.8	86.7		
Left	0.8	87.5		
Lemon	0.8	88.3		
Lenny	0.8	89.2		
Leonard	0.8	90.0		
Life	0.8	90.8		
Lily	0.8	91.7		
Lim	0.8	92.5		
Lincoln	0.8	93.3		
Lindsey	0.8	94.2		
Linus	0.8	95.0		
Lion	0.8	95.8		
Liz	0.8	96.7		
Lois	0.8	97.5		
Lola	0.8	98.3		
Lorraine	0.8	99.2	St Dev	Count
L: None	0.8	100.0	7.0	33

Table 25. Frequency distribution for M (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Mary	42.7	42.7		
Mike	8.7	51.5		
man	6.8	58.3		
money	4.9	63.1		
Mom	3.9	67.0		
mother	3.9	70.9		
moon	2.9	73.8		
Mario	1.9	75.7		
meal	1.9	77.7		
mountain	1.9	79.6		
mouse	1.9	81.6	*	11
mall	1.0	82.5		
manage	1.0	83.5		
maple	1.0	84.5		
marble	1.0	85.4		
Martha	1.0	86.4		
match	1.0	87.4		
mean	1.0	88.3		
memory	1.0	89.3		
Michael	1.0	90.3		
microbe	1.0	91.3		
middel	1.0	92.2		
milk	1.0	93.2		
millenium	1.0	94.2		
mine	1.0	95.1		
miserable	1.0	96.1		
modern	1.0	97.1		
Monday	1.0	98.1		
more	1.0	99.0	St Dev	Count
movie	1.0	100.0	7.7	30

Table 26. Frequency distribution for M (first names)

Name	Percent	Cumulative	Cutoff	Count
Mary	46.1	46.1		
Mike	10.2	56.3		
Michael	9.4	65.6		
Marc	6.3	71.9		
Mandy	1.6	73.4		
Maria	1.6	75.0		
Matthew	1.6	76.6		
Monica	1.6	78.1	*	8
Mako	0.8	78.9		
Malcom	0.8	79.7		
Man	0.8	80.5		
Mana	0.8	81.3		
Manor	0.8	82.0		
Marcel	0.8	82.8		
Margaret	0.8	83.6		
Matt	0.8	84.4		
Melissa	0.8	85.2		
Messiah	0.8	85.9		
Michelle	0.8	86.7		
Mom	0.8	87.5		
Moon	0.8	88.3	St Dev	Count
M: None	11.7	100.0	9.9	22

Table 27. Frequency distribution for N (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Nancy	49.5	49.5		
nice	5.8	55.3		
no	5.8	61.2		
November	5.8	67.0		
name	2.9	69.9		
new	2.9	72.8		
noise	1.9	74.8		
north	1.9	76.7	*	8
nature	1.0	77.7		
neat	1.0	78.6		
negotiate	1.0	79.6		
Nellie	1.0	80.6		
net	1.0	81.6		
Newton	1.0	82.5		
nickle	1.0	83.5		
night	1.0	84.5		
nine	1.0	85.4		
node	1.0	86.4		
non	1.0	87.4		
none	1.0	88.3		
nonsense	1.0	89.3		
noodle	1.0	90.3		
nope	1.0	91.3		
Norway	1.0	92.2		
nose	1.0	93.2		
note	1.0	94.2		
noteworthy	1.0	95.1		
nothing	1.0	96.1		
notoriety	1.0	97.1		
noun	1.0	98.1		
nurse	1.0	99.0	St Dev	Count
nut	1.0	100.0	8.6	32

Table 28. Frequency distribution for N (first names)

Name	Percent	Cumulative	Cutoff	Count
Nancy	72.9	72.9		
Nathan	3.4	76.3		
Nick	3.4	79.7	*	3
Nicole	1.7	81.4		
November	1.7	83.1		
Name	0.8	83.9		
Nan	0.8	84.7		
Napoleon	0.8	85.6		
Nate	0.8	86.4		
Ned	0.8	87.3		
Neil	0.8	88.1		
Nellie	0.8	89.0		
Nelson	0.8	89.8		
Neophite	0.8	90.7		
Nicholas	0.8	91.5		
Night	0.8	92.4		
Nina	0.8	93.2		
Noemie	0.8	94.1		
Normal	0.8	94.9		
Norman	0.8	95.8		
Normand	0.8	96.6		
Nupur	0.8	97.5	St Dev	Count
N: None	2.5	100.0	15.0	23

Table 29. Frequency distribution for O (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
open	21.4	21.4		
Oscar	18.4	39.8		
orange	12.6	52.4		
ocean	2.9	55.3		
octopus	2.9	58.3		
opera	2.9	61.2		
Oregon	2.9	64.1		
Ohio	1.9	66.0		
old	1.9	68.0		
operator	1.9	69.9		
other	1.9	71.8		
owl	1.9	73.8		
ox	1.9	75.7	*	13
object	1.0	76.7		
oboe	1.0	77.7		
ochre	1.0	78.6		
October	1.0	79.6		
omega	1.0	80.6		
often	1.0	81.6		
okay	1.0	82.5		
olive	1.0	83.5		
only	1.0	84.5		
onto	1.0	85.4		
opal	1.0	86.4		
operate	1.0	87.4		
opportunity	1.0	88.3		
or	1.0	89.3		
orchestrate	1.0	90.3		
orenthal	1.0	91.3		
oriental	1.0	92.2		
oscilator	1.0	93.2		
ostrich	1.0	94.2		
Otis	1.0	95.1		
otter	1.0	96.1		
out	1.0	97.1		
out there	1.0	98.1		
outside	1.0	99.0	St Dev	Count
over	1.0	100.0	4.6	38

Table 30. Frequency distribution for O (first names)

Name	Percent	Cumulative	Cutoff	Count
Oscar	38.1	38.1		
Oliver	6.8	44.9		
Oprah	6.8	51.7		
Omar	4.2	55.9		
Opal	3.4	59.3		
Open	3.4	62.7		
Owen	3.4	66.1		
Olga	1.7	67.8		
Olivia	1.7	69.5		
Ophelia	1.7	71.2		
Orlando	1.7	72.9		
Othello	1.7	74.6		
Otto	1.7	76.3	*	13
Ofeila	0.8	77.1		
Okay	0.8	78.0		
Olie	0.8	78.8		
Olive	0.8	79.7		
Oma	0.8	80.5		
Onyx	0.8	81.4		
Operator	0.8	82.2		
Opie	0.8	83.1		
Opposum	0.8	83.9		
Orpheus	0.8	84.7		
Osmond	0.8	85.6		
Otis	0.8	86.4		
Oyang	0.8	87.3	St Dev	Count
O: None	12.7	100.0	7.4	27

Table 31. Frequency distribution for P (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Paul	23.5	23.5		
peter	13.7	37.3		
Papa	5.9	43.1		
people	3.9	47.1		
play	3.9	51.0		
part	2.0	52.9		
party	2.0	54.9		
pen	2.0	56.9		
pet	2.0	58.8		
pig	2.0	60.8		
pirate	2.0	62.7		
pretty	2.0	64.7		
puppy	2.0	66.7		
push	2.0	68.6	*	14

Word	Percent	Cumulative	Cutoff	Count
package	1.0	1.0		
page	1.0	2.0		
pan	1.0	2.9		
paranoid	1.0	3.9		
park	1.0	4.9		
Pat	1.0	5.9		
Patrick	1.0	6.9		
peach	1.0	7.8		
pencil	1.0	8.8		
penny	1.0	9.8		
performance	1.0	10.8		
pewter	1.0	11.8		
pineapple	1.0	12.7		
pink	1.0	13.7		
pita	1.0	14.7		
plane	1.0	15.7		
planning	1.0	16.7		
pleasant	1.0	17.6		
please	1.0	18.6		
plum	1.0	19.6		
pod	1.0	20.6		
pony	1.0	21.6		
popcorn	1.0	22.5		
port	1.0	23.5		
possible	1.0	24.5		
potato	1.0	25.5		
power	1.0	26.5		
progression	1.0	27.5		
project	1.0	28.4		
proud	1.0	29.4		
purchase	1.0	30.4	St Dev	Count
purpose	1.0	31.4	0.0	33

Table 32. Frequency distribution for P (first names)

Name	Percent	Cumulative	Cutoff	Count
Paul	40.0	40.0		
Peter	35.8	75.8		
Papa	2.5	78.3		
Pat	2.5	80.8		
Patty	2.5	83.3		
Penny	2.5	85.8	*	6
Pete	1.7	87.5		
Philip	1.7	89.2		
Pauline	0.8	90.0		
Pear	0.8	90.8		
Pearl	0.8	91.7		
Penelope	0.8	92.5		
Pepito	0.8	93.3		
Percy	0.8	94.2		
Petra	0.8	95.0		
Ping	0.8	95.8		
Prema	0.8	96.7		
Print	0.8	97.5		
Priscilla	0.8	98.3		
Private	0.8	99.2	St Dev	Count
Purple	0.8	100.0	11.1	21

Table 33. Frequency distribution for *Q* (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
queen	31.1	31.1		
quick	11.7	42.7		
Quebec	8.7	51.5		
quiet	6.8	58.3		
quit	6.8	65.0		
question	5.8	70.9		
query	3.9	74.8		
quilt	3.9	78.6	*	8
quality	2.9	81.6		
quarter	2.9	84.5		
quack	1.9	86.4		
quest	1.9	88.3		
Quincy	1.9	90.3		
quite	1.9	92.2		
quiz	1.9	94.2		
quaint	1.0	95.1		
quarterly	1.0	96.1		
queer	1.0	97.1		
queue	1.0	98.1		
quirk	1.0	99.0	St Dev	Count
quota	1.0	100.0	6.7	21

Table 34. Frequency distribution for Q (first names)

Name	Percent	Cumulative	Cutoff	Count
Quency	13.6	13.6		
Quenten	13.6	27.1		
Queen	8.5	35.6		
Quin	5.1	40.7		
Qin	2.5	43.2		
Quebec	2.5	45.8		
Queenie	2.5	48.3		
Quick	2.5	50.8		
Qing	1.7	52.5		
Quail	1.7	54.2		
Quint	1.7	55.9		
Qamar	0.8	56.8		
Quack	0.8	57.6		
Qualls	0.8	58.5		
Quan	0.8	59.3		
Quark	0.8	60.2		
Quasar	0.8	61.0		
Quasimoto	0.8	61.9		
Quatli	0.8	62.7		
Quella	0.8	63.6		
Question	0.8	64.4		
Queta	0.8	65.3		
Quetzlquote	0.8	66.1		
Quilvio	0.8	66.9		
Quinones	0.8	67.8		
Quito	0.8	68.6		
Quixote	0.8	69.5		
Quotient	0.8	70.3	*	28
			St Dev	Count
Q: None	29.7	100.0	6.1	29

Table 35. Frequency distribution for R (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Robert	17.5	17.5		
rabbit	7.8	25.2		
Romeo	6.8	32.0		
Richard	5.8	37.9		
run	5.8	43.7		
Roger	4.9	48.5		
ralph	3.9	52.4		
ready	2.9	55.3		
red	2.9	58.3		
rose	2.9	61.2		
radio	1.9	63.1		
Randy	1.9	65.0		
rock	1.9	67.0		
rug	1.9	68.9		
rush	1.9	70.9	*	15
raccoon	1.0	71.8		
railroad	1.0	72.8		
rain	1.0	73.8		
rainbow	1.0	74.8		
rap	1.0	75.7		
rat	1.0	76.7		
raw	1.0	77.7		
real	1.0	78.6		
reality	1.0	79.6		
remarketer	1.0	80.6		
repeat	1.0	81.6		
replenishment	1.0	82.5		
rerun	1.0	83.5		
reset	1.0	84.5		
respect	1.0	85.4		
restaurant	1.0	86.4		
results	1.0	87.4		
Rex	1.0	88.3		
rhythm	1.0	89.3		
rich	1.0	90.3		
Rick	1.0	91.3		
rickety	1.0	92.2		
riddle	1.0	93.2		
right	1.0	94.2		
road	1.0	95.1		
robot	1.0	96.1		
root	1.0	97.1		
round	1.0	98.1		
Roxanne	1.0	99.0	St Dev	Count
royal	1.0	100.0	2.9	45

Table 36. Frequency distribution for R (first names)

Name	Percent	Cumulative	Cutoff	Count
Robert	38.3	38.3		
Roger	9.2	47.5		
Richard	8.3	55.8		
Rae	5.0	60.8		
Ralph	3.3	64.2		
Romeo	3.3	67.5		
Rick	2.5	70.0		
Rob	2.5	72.5		
Rachel	1.7	74.2		
Randy	1.7	75.8		
Rebecca	1.7	77.5		
Rene	1.7	79.2		
Ron	1.7	80.8		
Ronald	1.7	82.5		
Ruth	1.7	84.2	*	15
Racecar	0.8	85.0		
Rain	0.8	85.8		
Ran	0.8	86.7		
Randal	0.8	87.5		
Random	0.8	88.3		
Ravi	0.8	89.2		
Raymond	0.8	90.0		
Regina	0.8	90.8		
Right	0.8	91.7		
Rita	0.8	92.5		
Roberta	0.8	93.3		
Robin	0.8	94.2		
Rosco	0.8	95.0		
Rose	0.8	95.8		
Roy	0.8	96.7		
Rudolph	0.8	97.5		
Rueben	0.8	98.3		
Rule	0.8	99.2	St Dev	Count
Run	0.8	100.0	6.6	34

Table 37. Frequency distribution for S (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Sam	37.5	37.5		
snake	3.8	41.3		
star	3.8	45.2		
Sally	2.9	48.1		
start	2.9	51.0		
stop	2.9	53.8		
Seattle	1.9	55.8		
short	1.9	57.7		
Sierra	1.9	59.6		
stone	1.9	61.5		
sun	1.9	63.5	*	11

Word	Percent	Cumulative	Cutoff	Count
same	1.0	64.4		
Sammy	1.0	65.4		
Samuel	1.0	66.3		
sand	1.0	67.3		
Santa	1.0	68.3		
Saturn	1.0	69.2		
save	1.0	70.2		
schizophrenic	1.0	71.2		
sebra	1.0	72.1		
send	1.0	73.1		
sentence	1.0	74.0		
September	1.0	75.0		
service	1.0	76.0		
set	1.0	76.9		
sexy	1.0	77.9		
Sharon	1.0	78.8		
simple	1.0	79.8		
sister	1.0	80.8		
smart	1.0	81.7		
soda	1.0	82.7		
solo	1.0	83.7		
something	1.0	84.6		
song	1.0	85.6		
spaghetti	1.0	86.5		
Spike	1.0	87.5		
spring	1.0	88.5		
Stephen	1.0	89.4		
Steve	1.0	90.4		
still	1.0	91.3		
store	1.0	92.3		
strategy	1.0	93.3		
strong	1.0	94.2		
sugar	1.0	95.2		
summer	1.0	96.2		
super	1.0	97.1		
superficial	1.0	98.1		
Susan	1.0	99.0	St Dev	Count
Suzy	1.0	100.0	10.6	11

Table 38. Frequency distribution for S (first names)

Name	Percent	Cumulative	Cutoff	Count
Sam	49.2	49.2		
Sally	8.3	57.5		
Steve	6.7	64.2		
Sara	5.8	70.0		
Steven	5.0	75.0		
Susan	4.2	79.2	*	6
Sierra	2.5	81.7		
Samuel	1.7	83.3		
Sandy	1.7	85.0		
Smith	1.7	86.7		
Stacy	1.7	88.3		
Samantha	0.8	89.2		
Saswat	0.8	90.0		
Scott	0.8	90.8		
Sharon	0.8	91.7		
Shawn	0.8	92.5		
Shijun	0.8	93.3		
Shirley	0.8	94.2		
Sister	0.8	95.0		
Snake	0.8	95.8		
Sophie	0.8	96.7		
Stan	0.8	97.5		
Stanley	0.8	98.3		
Start	0.8	99.2	St Dev	Count
Sue	0.8	100.0	9.6	25

Table 39. Frequency distribution for T (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
Tom	35.3	35.3		
tango	6.9	42.2		
Thomas	4.9	47.1		
tree	4.9	52.0		
top	3.9	55.9		
toy	3.9	59.8		
time	2.9	62.7		
table	2.0	64.7		
talk	2.0	66.7		
tall	2.0	68.6		
tangle	2.0	70.6		
Tony	2.0	72.5	*	12
tablet	1.0	73.5		
taco	1.0	74.5		
tambourine	1.0	75.5		
teach	1.0	76.5		
teapot	1.0	77.5		
technology	1.0	78.4		
Ted	1.0	79.4		
template	1.0	80.4		
temple	1.0	81.4		
ten	1.0	82.4		
terrible	1.0	83.3		
terrific	1.0	84.3		
theoretical	1.0	85.3		
there	1.0	86.3		
tie	1.0	87.3		
tiger	1.0	88.2		
tommorw	1.0	89.2		
totally	1.0	90.2		
tough	1.0	91.2		
train	1.0	92.2		
trembling	1.0	93.1		
trick	1.0	94.1		
trouble	1.0	95.1		
truck	1.0	96.1		
tuba	1.0	97.1		
tunnel	1.0	98.0		
turtle	1.0	99.0	St Dev	Count
type	1.0	100.0	5.5	40

Table 40. Frequency distribution for T (first names)

Name	Percent	Cumulative	Cutoff	Count
Tom	55.0	55.0		
Thomas	6.7	61.7		
Tony	5.0	66.7		
Tango	4.2	70.8		
Tim	4.2	75.0		
Ted	2.5	77.5		
Terri	2.5	80.0	*	7
Tammy	1.7	81.7		
Tiger	1.7	83.3		
Tina	1.7	85.0		
Tomas	1.7	86.7		
Tal	0.8	87.5		
Tank	0.8	88.3		
Tara	0.8	89.2		
Tarence	0.8	90.0		
Teddy	0.8	90.8		
Teresa	0.8	91.7		
Teto	0.8	92.5		
Text	0.8	93.3		
Tia	0.8	94.2		
Tieying	0.8	95.0		
Todd	0.8	95.8		
Tommy	0.8	96.7		
Tracy	0.8	97.5		
Tube	0.8	98.3	St Dev	Count
T: None	1.7	100.0	10.5	26

Table 41. Frequency distribution for U (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
under	20.0	20.0		
uncle	14.7	34.7		
ugly	8.0	42.7		
union	6.7	49.3		
unicorn	5.3	54.7		
uniform	5.3	60.0		
use	5.3	65.3		
universe	4.0	69.3		
up	4.0	73.3		
user	4.0	77.3		
useful	2.7	80.0		
usual	2.7	82.7	*	12
ultimatum	1.3	84.0		
undefined	1.3	85.3		
underscore	1.3	86.7		
unforgettable	1.3	88.0		
unicode	1.3	89.3		
unique	1.3	90.7		
unit	1.3	92.0		
unite	1.3	93.3		
united	1.3	94.7		
universal	1.3	96.0		
unsettled	1.3	97.3		
unusual	1.3	98.7	St Dev	Count
Utah	1.3	100.0	4.5	25

Table 42. Frequency distribution for U (first names)

Name	Percent	Cumulative	Cutoff	Count
Ursula	20.5	20.5		
Ulysses	10.3	30.8		
Uncle	6.0	36.8		
Uma	5.1	41.9		
Under	5.1	47.0		
Umbrella	2.6	49.6		
Uniform	2.6	52.1		
Unis	1.7	53.8		
Uber	0.9	54.7		
Ugene	0.9	55.6		
Ujwala	0.9	56.4		
Ulric	0.9	57.3		
Ulyse	0.9	58.1		
Ungar	0.9	59.0		
Unice	0.9	59.8		
Unicorn	0.9	60.7		
Union	0.9	61.5		
University	0.9	62.4		
Unna	0.9	63.2		
Upchurch	0.9	64.1		
Uranium	0.9	65.0		
Uri	0.9	65.8	*	22
			St Dev	Count
U: None	34.2	100.0	7.9	23

Table 43. Frequency distribution for V (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
vertical	49.0	49.0		
victory	16.7	65.7		
violin	4.9	70.6		
very	3.9	74.5		
voice	2.9	77.5	*	5
velvet	2.0	79.4		
verify	2.0	81.4		
version	2.0	83.3		
vacant	1.0	84.3		
vacuous	1.0	85.3		
valve	1.0	86.3		
vector	1.0	87.3		
vegetable	1.0	88.2		
Venus	1.0	89.2		
verbose	1.0	90.2		
vertical	1.0	91.2		
vest	1.0	92.2		
vine	1.0	93.1		
violet	1.0	94.1		
virtue	1.0	95.1		
virus	1.0	96.1		
void	1.0	97.1		
volleyball	1.0	98.0		
volt	1.0	99.0	St Dev	Count
Volvo	1.0	100.0	9.9	25

Table 44. Frequency distribution for V (first names)

Name	Percent	Cumulative	Cutoff	Count
Victor	58.3	58.3		
Vickey	6.7	65.0		
Victoria	4.2	69.2		
Vincent	4.2	73.3		
Vera	3.3	76.7	*	5
Valerie	2.5	79.2		
Veronica	2.5	81.7		
Victory	2.5	84.2		
Vivian	2.5	86.7		
Very	1.7	88.3		
Vic	1.7	90.0		
Val	0.8	90.8		
Venetia	0.8	91.7		
Venus	0.8	92.5		
Vernecia	0.8	93.3		
Veronique	0.8	94.2		
Vessey	0.8	95.0		
Vicrot	0.8	95.8		
Vince	0.8	96.7		
Vinny	0.8	97.5		
Vito	0.8	98.3		
Vivtor	0.8	99.2	St Dev	Count
Volleyball	0.8	100.0	11.9	23

Table 45. Frequency distribution for W (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
water	9.0	9.0		
William	9.0	18.0		
whiskey	8.0	26.0		
white	5.0	31.0		
wally	4.0	35.0		
Walter	4.0	39.0		
Washington	4.0	43.0		
world	4.0	47.0		
why	3.0	50.0		
window	3.0	53.0		
wait	2.0	55.0		
wall	2.0	57.0		
walrus	2.0	59.0		
west	2.0	61.0		
wet	2.0	63.0		
whale	2.0	65.0		
wish	2.0	67.0		
woman	2.0	69.0	*	18

Word	Percent	Cumulative	Cutoff	Count
Waldo	1.0	70.0		
walk	1.0	71.0		
warm	1.0	72.0		
warren	1.0	73.0		
wax	1.0	74.0		
Wayne	1.0	75.0		
weather	1.0	76.0		
web	1.0	77.0		
Wednesday	1.0	78.0		
well	1.0	79.0		
well-meaning	1.0	80.0		
when	1.0	81.0		
wholesale	1.0	82.0		
willow	1.0	83.0		
Willy	1.0	84.0		
win	1.0	85.0		
winter	1.0	86.0		
wire	1.0	87.0		
wise	1.0	88.0		
with	1.0	89.0		
wolf	1.0	90.0		
women	1.0	91.0		
wonder	1.0	92.0		
wood	1.0	93.0		
woods	1.0	94.0		
word	1.0	95.0		
work	1.0	96.0		
worthy	1.0	97.0		
wounded	1.0	98.0		
wrong	1.0	99.0	St Dev	Count
Wyoming	1.0	100.0	2.4	18

Table 46. Frequency distribution for W (first names)

Name	Percent	Cumulative	Cutoff	Count
William	42.0	42.0		
Walter	13.4	55.5		
Wayne	5.9	61.3		
Wendy	5.9	67.2		
Wally	5.0	72.3		
Wanda	3.4	75.6		
Will	2.5	78.2		
Win	2.5	80.7	*	8
Walt	1.7	82.4		
Whiskey	1.7	84.0		
Wade	0.8	84.9		
Wallace	0.8	85.7		
Walrus	0.8	86.6		
Watch	0.8	87.4		
Watts	0.8	88.2		
Weizhe	0.8	89.1		
Wenzhi	0.8	89.9		
Wesley	0.8	90.8		
Whitney	0.8	91.6		
Widget	0.8	92.4		
Wilbur	0.8	93.3		
Williab	0.8	94.1		
Wilma	0.8	95.0		
Witch	0.8	95.8		
Witney	0.8	96.6		
Wolfram	0.8	97.5		
Wonderful	0.8	98.3	St Dev	Count
W: None	1.7	100.0	8.0	28

Table 47. Frequency distribution for X (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
x-ray	64.6	64.6		
xylophone	16.7	81.3	*	2
Xerox	11.5	92.7		
Xena	2.1	94.8		
Xanadu	1.0	95.8		
Xavier	1.0	96.9		
xenophobe	1.0	97.9		
xoom	1.0	99.0	St Dev	Count
xtricated	1.0	100.0	20.9	9

Table 48. Frequency distribution for X (first names)

Name	Percent	Cumulative	Cutoff	Count
Xavier	31.1	31.1		
Xray	13.4	44.5		
Xena	10.1	54.6		
Xenon	1.7	56.3		
Xerxes	1.7	58.0		
Xantipe	0.8	58.8		
Xerox	0.8	59.7		
Xia	0.8	60.5		
Xian	0.8	61.3		
Ximena	0.8	62.2		
Xingyu	0.8	63.0		
Xmas	0.8	63.9		
X-Rat	0.8	64.7		
Xylophone	0.8	65.5	*	14
			St Dev	Count
X: None	34.5	100.0	8.6	15

Table 49. Frequency distribution for Y (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
yellow	45.5	45.5		
yes	13.1	58.6		
Yankee	10.1	68.7		
young	6.1	74.7		
yoyo	5.1	79.8	*	5
yak	4.0	83.8		
yestarday	3.0	86.9		
you	3.0	89.9		
yell	2.0	91.9		
yard	1.0	92.9		
yatch	1.0	93.9		
yelping	1.0	94.9		
yeoman	1.0	96.0		
yield	1.0	97.0		
yolk	1.0	98.0		
your	1.0	99.0	St Dev	Count
yours	1.0	100.0	10.8	17

Table 50. Frequency distribution for Y (first names)

Name	Percent	Cumulative	Cutoff	Count
Yolanda	30.5	30.5		
Yello	6.8	37.3		
Yvonne	5.9	43.2		
Yankee	4.2	47.5		
Young	2.5	50.0		
Yvette	2.5	52.5		
Yen	1.7	54.2		
Yes	1.7	55.9		
Yale	0.8	56.8		
Yan	0.8	57.6		
Yancey	0.8	58.5		
Yandow	0.8	59.3		
Yang	0.8	60.2		
Yasar	0.8	61.0		
Year	0.8	61.9		
Yelena	0.8	62.7		
Yemen	0.8	63.6		
Yesterday	0.8	64.4		
Yihong	0.8	65.3		
Yin	0.8	66.1		
Yoddle	0.8	66.9		
Yolinda	0.8	67.8		
Yongan	0.8	68.6		
Yonnie	0.8	69.5		
Yorke	0.8	70.3		
Yoshi	0.8	71.2		
Youngster	0.8	72.0		
Yun	0.8	72.9	*	28
			St Dev	Count
Y: None	27.1		7.2	29

Table 51. Frequency distribution for Z (unconstrained)

Word	Percent	Cumulative	Cutoff	Count
zebra	75.2	75.2		
zoo	5.9	81.2	*	2
zero	5.0	86.1		
Zulu	5.0	91.1		
Zorro	3.0	94.1		
zip	1.0	95.0		
zipped	1.0	96.0		
zipper	1.0	97.0		
Zippo	1.0	98.0		
zoom	1.0	99.0	St Dev	Count
zygote	1.0	100.0	22.0	11

Table 52. Frequency distribution for Z (first names)

Name	Percent	Cumulative	Cutoff	Count
Zebra	18.4	18.4		
Zach	9.6	28.1		
Zachary	7.9	36.0		
Zelda	7.9	43.9		
Zoe	6.1	50.0		
Zeik	2.6	52.6		
Zane	1.8	54.4		
Zed	1.8	56.1		
Zoey	1.8	57.9		
Zorro	1.8	59.6		
Zulu	1.8	61.4		
Zebert	0.9	62.3		
Zebulun	0.9	63.2		
Zefferino	0.9	64.0		
Zeier	0.9	64.9		
Zell	0.9	65.8		
Zella	0.9	66.7		
Zenath	0.9	67.5		
Zeno	0.9	68.4		
Zerba	0.9	69.3		
Zernia	0.9	70.2		
Zeus	0.9	71.1		
Zevon	0.9	71.9		
Zhang	0.9	72.8		
Zhao	0.9	73.7		
Zhixin	0.9	74.6		
Zianna	0.9	75.4		
Zike	0.9	76.3		
Zita	0.9	77.2		
Zorba	0.9	78.1		
Zuber	0.9	78.9		
Zuzy	0.9	79.8	*	32
			St Dev	Count
Z: None	20.2		4.8	33

Correlation Analyses

Tables 53 and 54 show the correlations among the 80% cutoff points, standard deviations, and total word counts for both word distributions. In both distributions the cutoff points and total words counts had a positive correlation, and both of these variables had negative correlations with the standard deviations.

Table 53. Correlations Among Variables in the Unconstrained Distribution

Var 1	Var 2	Correlation	p
Cutoff	Std Dev	-0.87	0.00000001
Cutoff	Count	0.84	0.00000001
Std Dev	Count	-0.84	0.00000001

Table 54. Correlations Among Variables in the Names Distribution

Var 1	Var 2	Correlation	p
Cutoff	Std Dev	-0.64	0.0004
Cutoff	Count	0.50	0.01
Std Dev	Count	-0.80	0.000001

These correlations show the interesting (though perhaps not surprising) fact that the more words produced for a letter (which increases the 80% cutoff point and the total word count), the smaller the standard deviation.

Comparison of Unconstrained and Name Distributions

Table 55 shows a number of comparisons between the distributions. A set of *t*-tests (treating letters as subjects) indicated that there was no significant difference between the distributions for the variables of 80% cutoff point, standard deviation, and total word counts. There was a significant difference between the distributions for the variable of Nones – the percentage of responses for which respondents indicated that they couldn't think of a word for a given letter. This never occurred in the unconstrained distribution, but happened for a number of letters in the names distribution (most notably, Q, U, X, Y and Z).

Table 55. Comparisons Between Unconstrained and Name Distributions

Letter	CutoffN	StdDevN	CountN	CutoffW	StdDevW	CountW	CutoffD	StdDevD	CountD	Nones	
A	15	2.6	36	3	16.4	16	12	-13.81	20	0.0	
B	10	6.1	29	14	7.6	28	-4	-1.43	1	0.0	
C	8	6.0	32	3	10.6	23	5	-4.64	9	2.3	
D	7	9.0	25	2	16.8	15	5	-7.81	10	1.7	
E	19	4.7	39	11	4.8	32	8	-0.16	7	3.4	
F	3	16.6	14	11	6.8	31	-8	9.80	-17	7.1	
G	6	9.5	26	9	5.3	29	-3	4.14	-3	0.8	
H	10	8.0	23	11	4.0	27	-1	4.00	-4	1.7	
I	35	3.3	36	19	2.8	43	16	0.55	-7	12.0	
J	7	6.1	29	17	3.1	42	-10	2.93	-13	0.0	
K	14	4.3	29	17	4.4	32	-3	-0.09	-3	3.8	
L	11	7.0	33	17	3.8	39	-6	3.18	-6	0.8	
M	8	9.9	22	11	7.7	30	-3	2.22	-8	11.7	
N	3	15.0	23	8	8.6	32	-5	6.37	-9	2.5	
O	13	7.4	27	13	4.6	38	0	2.81	-11	12.7	
P	6	11.1	21	14	3.8	46	-8	7.25	-25	0.0	
Q	28	6.1	29	8	6.7	21	20	-0.59	8	29.7	
R	15	6.6	34	15	2.9	45	0	3.67	-11	0.0	
S	6	9.6	25	11	5.2	49	-5	4.42	-24	0.0	
T	7	10.5	26	12	5.5	40	-5	5.05	-14	1.7	
U	22	7.9	23	12	4.5	25	10	3.37	-2	34.2	
V	5	11.9	23	5	9.9	25	0	1.96	-2	0.0	
W	8	8.0	28	18	2.0	49	-10	6.00	-21	1.7	
X	14	8.6	15	2	20.9	9	12	-12.28	6	34.5	
Y	28	7.2	29	5	10.8	17	23	-3.58	12	27.1	
Z	32	4.8	33	2	22.0	11	30	-17.23	22	20.2	
Mean	13.08	7.99	27.27	10.38	7.75	30.54	2.69	0.23	-3.27	8.06	
StDev	9.03	3.28	6.06	5.35	5.53	11.58	10.74	6.63	12.48	11.38	
Count	26	26	26	26	26	26	26	26	26	26	
sem	1.77	0.64	1.19	1.05	1.08	2.27	2.11	1.30	2.45	2.23	
crit-t(90)	1.708	1.708	1.708	1.708	1.708	1.708	1.708	1.708	1.708	1.708	
delta(90)	3.02	1.10	2.03	1.79	1.85	3.88	3.60	2.22	4.18	3.81	
Upper	16.10	9.08	29.30	12.18	9.60	34.42	6.29	2.45	0.91	11.87	
Mean	13.08	7.99	27.27	10.38	7.75	30.54	2.69	0.23	-3.27	8.06	
Lower	10.05	6.89	25.24	8.59	5.90	26.66	-0.90	-1.99	-7.45	4.25	
							t	1.28	0.18	-1.34	3.61
							df	25	25	25	25
							p	0.21	0.86	0.19	0.001

Discussion

The hypothesis that asking for first names would produce more consistent responses than completely unconstrained responding did not hold. For some letters this was true, but for others it was not. Overall, the distributions had similar statistical properties with regard to 80% cutoff points, standard deviations, and total word counts (all measures of response consistency). However, the number of letters for which participants were not able to think of any appropriate response was greater for first names than for unconstrained words. For this reason, designers should avoid requiring users to provide first names when using words to perform voice spelling. One reasonable approach for limited voice spelling grammars would be to provide three candidates for each letter – the word from the military alphabet, the word with the greatest frequency for each letter from the unconstrained tables, and the word with the greatest frequency for each letter in the first names tables. Depending on the capacity of the system, it would be reasonable to include more of the words from each source to increase the odds of having a match to a user's choice of word when voice spelling.

References

Hartley, M. W., and Lewis, J. R. (2001). *Conditional probabilities for IBM voice browser recognition of letters of the alphabet* (Tech. Report 29.3421). West Palm Beach, FL: International Business Machines Corp.