ABSTRACT

SAMANTHA KNOLL
Data From the LAMSAS Project: How Language Varies in Systematic Ways
(Under the Direction of Dr. William A. Kretzschmar, Jr.)

According to Kretzschmar, the linguistics of linguistic structure can be differentiated from the linguistics of speech based on the idea of the linguistic continuum: "the continuously variable behavior of individual speakers" (Linguistics 52). In other words, language is always highly variable because people make choices (often unconsciously) about what they say or how they say it. Although living human languages are constantly shifting and are not static, speech is "not chaotic or unmanageable", but instead exhibits regularities across geographic and social variables (Kretzschmar, Linguistics 52). Linguistic features demonstrate a remarkable property, in that "there will be few realizations that occur very frequently, and a great many realizations that occur only infrequently", which, when graphed, appears as an asymptotic hyperbolic curve, or A-curve (Kretzschmar, Linguistics 83). I examine in detail twelve lexical items from existing LAMSAS survey data, and the frequency counts of these lexical items exhibit characteristics of the A-curve, even when different subsets of the entire population are examined in isolation. Kretzschmar describes complex systems as "open and not at equilibrium," and showing "self-organization and the emergence of order" (Dialectology). These systems are also made up of many components interacting with one another, and they exhibit non-linear distribution as well as scaling (Kretzschmar, Dialectology). Although human language exists in many forms, it varies systematically and conforms to the principles of the theory of complex systems.

INDEX WORDS: Linguistics, Language, Complex Systems, Chaos Theory, LAMSAS, Linguistic Atlas, A-Curve, University of Georgia

DATA FROM THE LAMSAS PROJECT:

HOW LANGUAGE VARIES IN SYSTEMATIC WAYS

by

SAMANTHA KORRINN KNOLL

A Thesis Submitted to the Honors Council of the University of Georgia

in Partial Fulfillment of the Requirements for the Degree

BACHELOR OF ARTS

in LINGUISTICS

with HIGHEST HONORS

Athens, Georgia

2009

DATA FROM THE LAMSAS PROJECT:

HOW LANGUAGE VARIES IN SYSTEMATIC WAYS

by

SAMANTHA KORRINN KNOLL

Approved:

Dr. William A. Kretzschmar, Jr. — April 3, 2009
Dr. William A. Kretzschmar, Jr. — Date
Faculty Research Mentor

Approved:

Dr. Paula Warrington — April 3, 2009
Dr. Paula Warrington — Date
Reader

Approved:

Dr. David S. Williams — May 8, 2009
Dr. David S. Williams — Date
Director, Honors Program, Foundation Fellows and
Center for Undergraduate Research Opportunities

Approved:

Dr. Pamela B. Kleiber — May 8, 2009
Dr. Pamela B. Kleiber — Date
Associate Director, Honors Program and
Center for Undergraduate Research Opportunities

ACKNOWLEDGEMENTS

This thesis would not have been possible without the guidance of Dr. William A. Kretzschmar, Jr., who supported me enthusiastically during the entire process of researching, writing, and revising. His passion for knowledge shows in everything he accomplishes, and I want to thank him for inspiring me and encouraging me to explore my own interests.

I would also like to thank Dr. Paula Warrington for being so gracious with her time and advice, and for being an exceptional professor and friend during my time at the University of Georgia.

Finally, I would be remiss if I did not offer a heartfelt “thank you” to Jordan Lynn, my amazingly supportive fiancé and best friend. He has spent countless hours listening to my ideas, offering suggestions, and providing strength for those moments when I began to feel overwhelmed.

To Dr. Kretzschmar, Dr. Warrington, Jordan, and the rest of my family and friends: thank you all for your support and guidance, and for allowing me to follow my dreams.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1
INTRODUCTION

The suggestion that language can be described in terms of complex systems theory is a significant, though fairly recent idea. Many modern academics view language from a structuralist perspective, which assumes that each language is based on a system of rules shared by all speakers of that language in roughly homogenous speech communities (Kretzschmar, Linguistics 10-11). Similarly, generative linguists suggest that any specific human language arises from a universal system of rules, also assuming the existence of homogenous speech communities (Linguistics, 11). While each of these approaches has something significant to offer to the study of linguistics, both have difficulty in explaining the vast amount of variation that exists within all human languages. In order to study language variation adequately, it is necessary to employ a theory that accounts for the fact that humans make choices (often unconsciously) about the language they use, and that homogenous speech communities are rare, if not non-existent. William A. Kretzschmar, Jr. has generated just such a theory and termed it the "linguistics of speech," stating that this approach "assumes that everybody and every group and every place, every situation is different" (Linguistics 12).

The linguistics of speech is distinguished from rule-based approaches to the study of language by four major assumptions: "1) the continuum of linguistic behavior, 2) extensive (really massive) variation in all features at all times, 3) importance of regional/social proximity to 'shared' linguistic production, 4) differential frequency as a key factor in linguistic production both in regional/social groups and in collections of text corpora" (Kretzschmar, Linguistics 8). These four assumptions form the basis of the theory of complex systems, which are "open and

not at equilibrium…, show[ing] self-organization and the emergence of order" (Kretzschmar, Dialectology). These systems are also made up of many components interacting with one another, and they exhibit non-linear distribution as well as scaling (Kretzschmar, Dialectology). Human language, in its many forms, varies systematically and conforms to these principles; therefore, it can be analyzed according to the theory of complex systems.

Each of these four assumptions will be examined in detail, using twelve lexical items from existing linguistic survey data. The American Linguistic Atlas Project ("Linguistic Atlas Projects") provides data from surveys conducted as early as 1930, and more specifically, the Linguistic Atlas of the Middle and South Atlantic States (LAMSAS) specifically references data from the geographic area of the original 13 colonies and includes the following states: New York, New Jersey, Pennsylvania, Virginia, West Virginia, Delaware, Maryland, District of Columbia, North Carolina, South Carolina, Georgia, and Florida.

This examination will show that such linguistic features, when graphed, always exhibits an A-curve, a "power law" distribution (Kretzschmar, "Distributional" 378), and characteristics of scaling. These features are significant because they provide a framework for analyzing how people perceive the language that constantly occurs around them: it is theorized that the A-curve is utilized subconsciously by speakers of a language to determine what is "normal" or "different" for their particular social or regional subgroup (Kretzschmar, "Neural" 339). Such perceptions may lead to the designation of a "dialect," which can only be a generalization because it "misrepresents the actual distribution of variants in the group—which in fact…always has an A-curve pattern" (Kretzschmar, "Neural" 341). Finally, it will be shown that the A-curve distribution for linguistic features is stable over time, which will have a profound impact on the way linguists study language change.

CHAPTER TWO
LITERATURE REVIEW

The idea that complex systems and A-curves can be used to describe linguistic phenomena has been examined in detail by only a few researchers, including Joan Bybee (Organization, Emergence), William A. Kretzschmar, Jr. ("Distributional," Dialectology, "Neural", Linguistics), and Lynne Cameron and Diane Larsen-Freeman ("Complex," "Research"). As early as 1921, though, Edward Sapir recognized the immense variability that exists in human language, stating:

> Language is variable. Two individuals of the same generation and locality, speaking precisely the same dialect and moving in the same social circles, are never absolutely at one in their speech habits. A minute investigation of the speech of each individual would reveal countless differences of detail…. In a sense they speak slightly divergent dialects of the same language rather than identically the same language. (147)

Ronald R. Butters suggests that because of this vast amount of variation, chaos theory is applicable to the study of languages and speech. He claims that the well-known concept of the butterfly effect is of "particular relevance" to linguistic study (205) and quotes Edgar Schneider as saying: "In language evolution, we have seemingly random, and insignificant, variation, which at some point becomes systematic, begins to spread, and typically will be subconsciously or consciously loaded with some socially signalling function" (Butters 205). According to Butters, "random chance can in itself produce patterned results if (a) it operates over a large enough set of events and/or (b) there is an interrelatedness among the events such that the output

of one instance can affect the next instance" (205). Language does indeed include a substantial number of speech events and these events are highly interactive because many utterances are spoken between two or more people; however, this does not necessarily mean that language neatly conforms to the principles of chaos theory, as Bybee, Kretzschmar, and Cameron and Larsen-Freeman reveal in their respective works.

In essays as early as the 1980s, Joan Bybee was examining language as a complex system, or an emergent structure, although the hypothesis was not at first a popular one. The rise of corpus linguistics in the 1990s and the ability to examine large amounts of data with computers were turning points in the discipline, allowing for the popularization of the hypothesis that "grammar comes about through the repeated adaptation of forms to live discourse" (Bybee and Hopper 2). The authors explain the contrasting hypothesis of structuralism and its origins:

> A legacy of the structural tradition in linguistics is the widespread acceptance of the premise that language structure is independent of language use. This premise is codified in a variety of theoretical distinctions, such as *langue* and *parole* (Saussure 1916) and competence and performance (Chomsky 1965). A further premise of this legacy is that the study of structure is a higher calling than the study of usage and is a potentially more promising avenue for uncovering the basic cognitive mechanisms that make human language possible. (Bybee and Hopper 1).

The tenants of structuralism, which have been widely accepted in linguistics since Saussure's work, divide grammar from discourse and thus separate structure from use. In contrast, Bybee explains, since the 1970s, a small number of linguists have viewed grammar as "arising from the patterns of language use in actual discourse… In this view, it is proposed that

grammar arises diachronically because of the commonly used discourse patterns that humans need to communicate" (Bybee 6). Such linguists, Bybee included, believe that detaching the structures of language from how they are used in actual discourse "removes a valuable source of explanation for why language has grammar and what form that grammar takes" (Bybee 6).

Additionally, Bybee develops the concept of language as a complex system, although she applies the theory metaphorically:

> In complex systems, a small number of mechanisms operate in real time and with repetition lead to the emergence of what appears to be an organized structure, such as a sand dune. However, we know that a sand dune is not fixed in time and space but is ever altering and becoming. So we see that language is also always in a process of becoming—creating, losing, and re-creating structures that are never absolutely fixed, allowing for continued variation and change" (Bybee 8).

Bybee and Hopper still follow the generative paradigm and are more concerned with examining linguistic structure rather than actual discourse, though they make a crucial point in the introduction to their compilation of essays about frequency of use and linguistic structure. The idea that language is emergent "constitutes a break with standard ideas about grammar that envisage it as a fixed synchronic system. It relativizes structure to speakers' actual experience with language, and sees structure as an on-going response to the pressure of discourse rather than as a pre-existent matrix" (Bybee and Hopper 3). While previous hypotheses about language have been focused on "the broader patterns of structure and the more abstract and generalized categories" (Bybee 6), Bybee was one of the first linguists to express the importance of examining individual words and expressions, which is precisely what my own study does.

In a paper presented at a 2008 "Methods in Dialectology" conference, Kretzschmar explains that several observable facts about language contrast with what one would expect to find if language were "chaotic" in the sense described above. Kauffman describes a chaotic system (or "equilibrium system," as he labels it) as closed and static, with no energy entering or leaving the system (Kretzschmar, Linguistics 146), while languages are observably "open and dynamic" systems: "New conversations and new writings among members of any speaking population occur continuously at unimaginable rates, surely a dynamic system. If these things no longer occur for a language, we call it 'dead'" (Kretzschmar, Dialectology 10). Additionally, Kretzschmar challenges the assertion of Butters and Schneider, that language is full of "random" variation, by observing that "sounds and words are not randomly distributed in speech interactions, but instead are associated in different ways with particular localities, particular social groups, and particular text types" (Dialectology 10). Finally, although chaotic systems exhibit cyclical behavior over time, complex systems "show non-linear distribution of units, as opposed to random or statistically normal distribution" (Kretzschmar, Dialectology 8-9), a point which will be illustrated with my own evidence from actual speech in a subsequent section.

Lynne Cameron and Diane Larsen-Freeman also claim that viewing language as a complex system accounts for the facts of human speech better than viewing it as a chaotic system. Unlike many modern academic views of language variation and change, the approach taken by Cameron and Larsen-Freeman is not concerned with devising rule systems or pathways for change; instead, these two researchers claim that language change results from the constant interactions between humans and the choices each person makes while employing language (Complex 231). Language, they believe, is "as much a process as a product, something in which one participates… Because language is a dynamic system, continuously changing, its potential

too is always being developed, and it is never fully realised" (Complex 231). This approach also accounts for the immense variation of human languages because of its assumption that "language is not a single homogeneous construct to be acquired; rather, in the complex systems view that sees language as resulting from use, the centrality of variation and speakers' choice of lexicogrammatical constructions within a social context is foregrounded" (Cameron and Larsen-Freeman, Complex 231). Complex systems are described as being "composed of elements or agents that interact in different ways. Their interactions lead to self-organization and the emergence of new patterns at different levels and timescales" (Cameron and Larsen-Freeman, Complex 227). In a later section, it will be shown that language exhibits both self-organization and scaling, which is crucial support for the theory of language as a complex system.

One of the best sources of linguistic data from the 20th century comes from the *Linguistic Atlas of the United States*. Hans Kurath was responsible for beginning this project, which divides the United States into ten regions based on differences in lexical items, pronunciations, and certain grammatical features. Nina Brown, part of the Center for Spatially Integrated Social Science, describes the techniques used by Kurath and his researchers in the early 1930s:

> [A] small team of linguists fanned out across the region interviewing at least two people in every county. Kurath gave the researchers explicit instructions about the types of informants who were considered appropriate for the project. In every town or city selected for the project at least two people would be chosen, one had to be "old-fashioned and unschooled," Kurath suggested a farmer or a farmer's wife, and the other should be "a member of the middle class who has had the benefit of a grade-school or high-school education" (Kurath 1949: v). The communities themselves were also carefully screened. Kurath placed a priority

> on towns that were early American settlements or could be directly linked to them through historical records. (Brown)

Although Kurath's methods would be considered faulty by today's research methodology standards, Kretzschmar explains that "Kurath's method may have introduced some bias, [but] it would be difficult to say that it biased his results in any particular direction… There is no evidence that…speakers were included or rejected for how they spoke" (Linguistics 99). He also makes the poignant observation that "to throw out the historical evidence of 1162 speakers just because survey research standards have changed would be silly" (Kretzschmar, Linguistics 98). Instead, an ongoing modern survey of the same magnitude should be conducted in order to align those results with what is already available to twenty-first century linguists.

Kurath published *A Word Geography of the Eastern United States* in 1949, in which he analyzes *Atlas* data in an attempt to establish a scientific history of the American vocabulary (Kurath vi). While Kurath's approach is more focused on an historical analysis of language *change*, some of his observations are significant for the current discussion of language *variation*, such as his claim that "Under the dominance of these urban centers [seaports along the East Coast] local expressions and pronunciations have been replaced in the countryside by new expressions and pronunciations radiating from them. We can observe this trend from local to regional usage most clearly in the Boston area, the Philadelphia area, and the Virginia Piedmont" (Kurath 2). Kurath explains the difference between speech in New England and the Southern States as occurring because of different social structures in the two locations. In the New England seaports, "social caste was much more pronounced," while in Southern "plantation country all elements of the population except the planters and the upper classes in the seaports

lived in marked isolation. This fact is strikingly reflected to this day in great local differences in the speech of the simple folk, both white and Negro" (Kurath 5-6).

Another reason Kurath's *Word Geography* is such a monumental work in American linguistics is because of its focus on everyday speech, rather than "literary language" or "cultivated speech" (Kurath 9). As the author explains:

> Students and teachers of English have focused their attention almost exclusively on the literary language and on cultivated speech—often enough without proper regard for the existing regional differences in the speech of the best educated. Folk speech has been dabbled in by scholars and by amateurs. But the speech of the large middle class has hardly been touched by trained linguists despite a lively popular interest in this subject. (9)

Kurath points out that in the 1930s folk words were being replaced by items derived from common or cultivated speech, or even literary language (49); this is an observation that will come into play in my own analysis of LAMSAS data. Finally, Kurath was interested in dividing the East Coast of the United States into separate speech areas based on his analysis of vocabulary, grammar, and pronunciation of survey items (Brown). My approach to the LAMSAS data focuses less on drawing dialect boundaries and more on the striking regularities that exist in language, even in its wide variety of forms along the Eastern Coast of the United States.

Using data from the Linguistic Atlas Projects, Kretzschmar conducted a study to determine whether this speech data conforms to the idea of language as a complex system. He observed three key features from the survey data he analyzed:

> First, we see that variation exists everywhere. In our survey data, no target item possesses an invariant response. Second, the variants of the linguistic features from our survey…are distributed according to what statisticians call a "power law." That is, if we count the frequency of occurrence of each realization (whether lexical or phonetic) for a target item, we find that there will be few realizations that occur very frequently, and a great many realizations that occur only infrequently. Finally, we have observed that variants of linguistic features are highly likely to occur in geographical clusters, that many variants are significantly associated with social characteristics, and that the same variant may be significantly associated with multiple sociocultural characteristics at the same time. ("Distributional" 378)

While the fact that variation exists in human speech was noted by Sapir in 1921 and many others since, the *amount* of difference continues to surprise linguists even today. For example, in the LAMSAS data analyzed by Kretzschmar, there are over 100 different responses for the target item *thunderstorms*, and even after deletion of inappropriate responses and the combination of inflected forms with base forms (a process called lemmatization), 73 different responses are still observed ("Distributional" 380). Even "stable referents" like *mantel* produced more than 30 different responses ("Distributional" 380).

The second and third observations made by Kretzschmar in "Distributional Foundations for a Theory of Language Change" are the focus of my own examination. The items examined in his study, when graphed according to frequency, all exhibit an asymptotic hyperbolic curve (A-curve), characterized by a small number of highly frequent responses and a much larger set of less-frequently-occurring responses (Kretzschmar, "Distributional" 386). The present concept of

the A-curve has its roots in Zipf's Law, which states that once words in a text are placed in descending order according to frequency "you [can] multiply the frequency times the rank, [and] you get a number that remains (roughly) constant for every word in the text" (Kretzschmar, "Distributional" 386). In the 1980s, Benoit Mandelbrot improved upon Zipf's basic idea with a more accurate formula, strengthening and reaffirming the fundamental pattern of distribution ("Distributional" 386). Kretzschmar's analysis of the LAMSAS survey data illustrates that the distributional pattern described by Zipf and Mandelbrot is not confined to texts, but also exists in speech, so that when the responses given for the target item *thunderstorm* are organized by frequency in descending order, the graph appears as an A-curve, as shown below (Kretzschmar, "Distributional" 388).

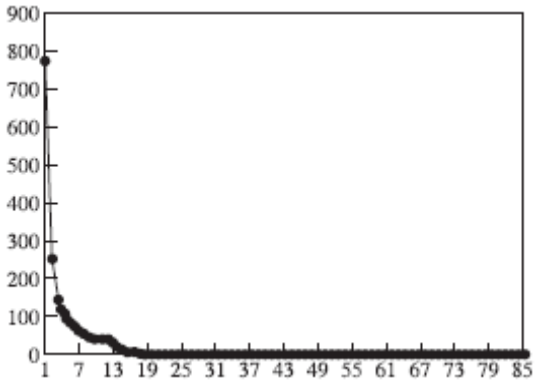


Figure 1: Frequency Graph for target item *thunderstorm* (Kretzschmar, "Distributional" 382).

The third of Kretzschmar's observations from LAMSAS data involves what he terms "non-random extralinguistic correlates" ("Distributional" 391). A 1996 study conducted with Schneider focused on two lexical items from LAMSAS data (*clearing up* and *cow lot)* and resulted in the observations that "a large number of the variants had statistically significant associations with one of eight independent variables (location, type, sex, age, race, community type, education, and occupation), and that the same variant could have significant associations

with several independent variables at the same time" (Kretzschmar, "Distributional" 391). Based on this, and other similar findings, it can be assumed that such variants are likely to be connected to more than one geographical and social factor at a time, and in quite an intricate fashion (Kretzschmar, "Distributional" 393). These findings also support the assumption that A-curves can be found in the frequency graphs for geographic and social subsets, an idea which will be examined in detail in my own analysis of LAMSAS data.

In *The Linguistics of Speech*, Kretzschmar establishes the basic foundations for his own approach to the study of language variation, called the "linguistics of speech," which is based on the idea that language exists as a complex system. He asserts that there are five basic requirements that language will satisfy if it indeed qualifies as a complex system: "1) speech is open and dynamic, thus not at equilibrium; 2) speech includes a very large number of interactive components/agents; 3) speech shows emergent order; 4) the distribution of units in speech is non-linear; 5) speech has the property of scaling" (151-152). The first two components have been dealt with above, and the remaining three will be examined in detail using graphical representations of frequency data from LAMSAS survey items. If this data satisfies the requirements of emergent order, non-linear distribution, and scaling, complex systems theory will have earned a place in English linguistics, not as a replacement for other theories, but as a new way of viewing linguistic data that may well provide fascinating insights about language variation and result in provocative consequences for the study of linguistics.

CHAPTER THREE
METHODS

The main source of data for this research comes from the Linguistic Atlas of the Middle and South Atlantic States, available online at http://us.english.uga.edu/cgi-bin/lapsite.fcgi/lamsas/. Many of the survey responses for lexical items and pronunciations have been digitized, and I chose lexical items to be analyzed from four broad categories or domains. These four categories were Agriculture/Land, Food, Home/Household Items, and Weather, chosen in the hopes that both women and men from upper and lower socioeconomic classes would have at least basic knowledge of all four domains, and because of a crucial observation made by Hans Kurath: "Regional and local expressions are most common in the vocabulary of the intimate everyday life of the home and the farm—not only among the simple folk and the middle class but also among the cultured" (9-10).

Once these four domains were established, I browsed the LAMSAS website and chose three lexical items from each domain. I browsed the list for any lexical item that fit into one of the four categories and decided upon three which had fewer than five "NA" ("not applicable") or "NR" ("no response") responses on pages 1, 5, 10, 15, and 20. The resulting 12 lexical items are listed below.

Agriculture/Land Domain: hog pen, meadow, swamp

Food Domain: cobbler, cornbread, pancakes

Home/Household Items Domain: andirons, hearth, pallet

Weather Domain: cloudburst, dry spell, steady drizzle

Once the lexical items were chosen, I downloaded the CSV (comma separated values) file from the LAMSAS website and converted it into spreadsheet format. I then sorted the lexical items alphabetically and deleted any inappropriate responses (see Appendix A for a list of deleted responses for each lexical item). All "NA" and "NR" responses were deleted, as well as any responses that seemed significantly out of place or anomalous. For example, the lexical item "pancakes" included 72 "NR" responses which were deleted, as well as "darkies called them," "my favorite," and 11 other such inappropriate responses.

After deleting such responses from the base spreadsheet, the data for each lexical item was organized into four different categories or subsets: all responses, male and female responses, three age ranges, and three sets of states. For the age groupings, the lexical items were sorted based on the ages of the respondents and then organized from youngest to oldest. In most cases, a few responses were not included because the age of the respondent was not recorded. The remaining responses were then divided into three equal groups, resulting in three age groups (the age group data for each lexical item is shown in Appendix C). Whenever the number of responses was not divisible by three, resulting in one or two extra responses, the first extra response was included in the first age group, and the second was included in the middle age group, since the LAMSAS data is somewhat biased towards older respondents.

For each lexical item, the state groupings were the same, despite an unequal distribution in the number of responses among the three groups. The groupings were determined based on the idea that state lines are not necessarily boundaries for different lexical items, an idea that is also held by Kretzschmar, who says that "variants for linguistic features just do not restrict themselves to neat areas" (Linguistics 69). Instead, the idea affecting the geographical groupings I created is that we as human beings "tend to talk like the people we are close to, whether the

closeness happens to be geographical or social" (Kretzschmar, "Distributional" 393). Thus, an attempt was made to organize the Middle and South Atlantic states into three groups based on fairly straight lines drawn across a map. The resulting state groups are as follows: Group 1: New York, New Jersey, Pennsylvania (henceforth abbreviated as NY/NJ/PA); Group 2: Maryland, District of Columbia, Delaware, Virginia, West Virginia (abbreviated as MD/DC/DE/VA/WV); and Group 3: North Carolina, South Carolina, Georgia, Florida (abbreviated as NC/SC/GA/FL). Often, Group 3 had significantly more responses than the other two groups due to the high number of respondents from South Carolina. The number of responses from each state per lexical item is included in Appendix D.

Once these four categories were established, the frequency counts of lexical items began. Inflected forms (typically plurals) were grouped with their non-inflected base forms, so that "marsh" and "marshes" are both counted under "marsh" in the frequency charts. In a few instances, it was also assumed that a certain possibly anomalous response was equivalent to a highly-occurring response, such as in the case of "brown bettu," which occurred only once in the data for the "cobbler" lexical item and was assimilated into the count for "brown betty" based on the assumption that "brown bettu" resulted from an error in transcription. Kretzschmar explains that such lemmatization does not significantly affect the outcome of the frequency analysis: "The large number of single occurrences does not go away in the set of responses without pluralizations and possibly inappropriate responses. The 'tail' of the curve is shortened and the basic distribution remains; this is true in our experience no matter how severely we lemmatize and otherwise restrict the response types" ("Distributional" 384). For the final frequency count, in instances with combined responses, the most common response is the one listed in the chart

(so that “pancakes” is shown, rather than “pancake,” which occurred less frequently). The combined responses for each lexical item are listed in Appendix B.

For each lexical item, a list of responses (after deletion and lemmatization) for each of the nine variables described above was determined and then organized by frequency in descending order. A line graph was then made from each frequency chart and, as will be described and explained below, each of the variables exhibited an A-curve, although the frequency of the lexical items in each chart varies. For these graphs, the variants of a lexical item are represented by the x-axis, and the y-axis represents the number of times each variant occurs. Each of the graphs has the same height and width proportions so that they can be compared to one another accurately. The “all responses” graphs for each lexical item are shown in Appendix E.

Every single variant that a respondent gives is of importance, but for the purposes of this study, only the top responses for each target lexical item are examined. Since there is a significant amount of difference in the number of variants given for each lexical item, the decision was made to examine only the top 10 responses for each subset in order to maintain some level of consistency between the lexical items and domains. Once the lists of responses were organized in descending order based on frequency, the list was truncated at the tenth response (or at the end of the list for those items with fewer than 10 unique responses). I have kept the data beyond the tenth response intact for the purposes of future studies, especially those on language change over time, but that data will not be used in this examination.

CHAPTER FOUR
RESULTS

For each of the following lexical items, 1,162 respondents were surveyed, although the number of total responses varies. For some of these items, a respondent may have provided more than one response, while for other items no response was provided by some of the respondents. These responses were typically elicited by the surveyors in conversation with the respondents.

Within the Agriculture/Land domain, the three target lexical items examined were "hog pen," "meadow," and "swamp[1]."

"Hog Pen" Data

For this target lexical item, there were a total of 1,554 responses and 74 different items. The distribution is that of a single highly-occurring item, a moderate number of items with more than one response, and 45 items for which just a single response was provided. The top 10 responses chart for the All Responses group is shown below.

Table 1: Top 10 Responses for "Hog Pen" – All Responses

hog pen	815
pig pen	272
pen	68
hog lot	64
pig sty	62
hog house	49
hog pasture	40
sty	36
fattening pen	17
floored pen	16

[1] See Appendix F for a complete collection of the "Top 10 Responses" charts for each lexical item.

For each subset of the respondent population for this lexical item, “hog pen” is the number one response, and typically by a wide margin—for example, the Male subset responded with “hog pen” 588 times, while the second most common item had only 158 responses. “Pig pen” was the second most common item for each subset except NC/SC/GA/FL: for this subset, “pen” was the second most common response, and “pig pen” was only the fourth most common. Once we reach the third most common item, the responses begin to vary more widely, with four unique items in the third most frequent position among the eight subsets: “pig sty” (Age Group 1, Females, NY/NJ/PA), “hog lot” (Age Group 2, Males, NC/SC/GA/FL), “pen” (Age Group 3), and “hog house” (MD/DC/DE/VA/WV). While “pen” does not occur until the fourth position or lower for each of the subsets except NC/SC/GA/FL and Age Group 3, it is the third most common item in the All Responses list. Almost all of the items in the fourth and fifth most frequent positions occurred for different subsets in third place, with the exception of “pig pen” (NC/SC/GA/FL), “hog pasture” (Age Group 3 and NC/SC/GA/FL), and “sty” (NY/NJ/PA). This pattern continues down the list, with the same few responses appearing in different orders for the different subsets. For example, “fattening pen” and “floored pen” typically alternate between 9th and 10th most frequent (these two items occur in the sixth and eighth positions, respectively, for the NC/SC/GA/FL subset). Finally, for this set of data, there are a few responses which appear in only one subset, including “hog sty” (Age Group 1), “pin” (Females), “pig house,” “pig yard,” and “pig stable” (NY/NJ/PA), and “hog bed,” “hog pound,” and “hog nest” (MD/DC/DE/VA/WV).

Most of the subsets had the same lexical items in their top 10 responses lists, although the exact placement of each response varies among the different subsets. Two of the geographic subsets show more variation than the other subsets: NY/NJ/PA and MD/DC/DE/VA/WV each

have three anomalous items in their top 10 responses. Among these outlying responses, all of those from the NY/NJ/PA subset use the word "pig," while the responses from MD/DC/DE/VA/WV employ the word "hog" instead. Both of these regions, however, have items in their top 10 lists that include "hog" and "pig," and the same is true for the NC/SC/GA/FL region, as well as all of the other subsets.

The All Responses graph[2] for the "hog pen" data illustrates an A-curve with Distribution Type A: a few items which occur quite frequently form the left-most part of the graph and a great number of single occurrences form the "tail" of the graph. A-curves with Distribution Type A are characterized by a rapid initial drop, followed by a gentler decline, and terminating with a long tail made up of items with only one or two responses.

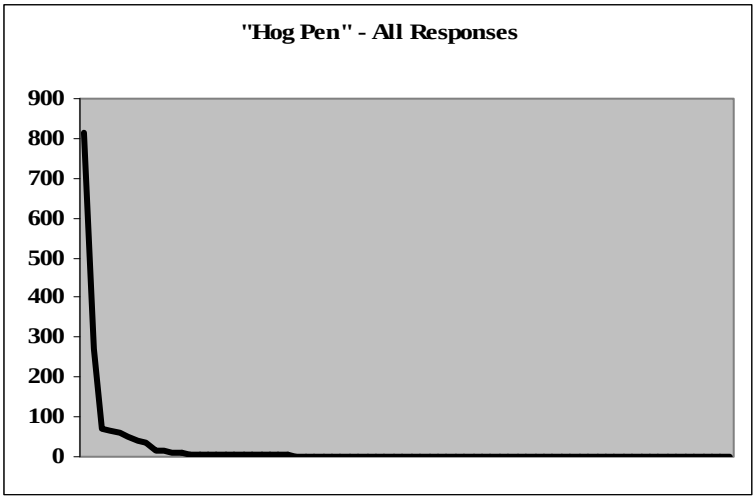


Figure 2: "Hog Pen" All Responses

Most of the "Hog Pen" subset graphs[3] look very similar to that of the All Responses graph, with dips and bumps in similar positions—in fact, the graph for Male responses looks practically identical to the All Responses graph.

[2] Each of the graphs included in this section and in Appendix E include *all* of the responses provided, not just the top 10 items.
[3] See Appendix E for a complete collection of the All Responses and subset graphs for each lexical item.

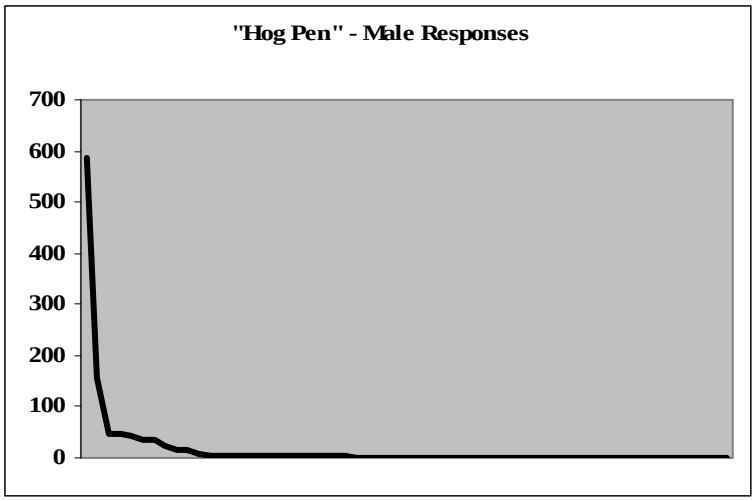


Figure 3: "Hog Pen" Male Responses

The graph for the NY/NJ/PA subset looks significantly different, although it still follows a general A-curve shape[4]: this difference can be explained by the relatively small number of unique responses in this subset (there are only 18, as compared to 58 for Males). Additionally, the tail of the graph begins later than the other graphs, due to the fact that among the 18 unique items, 12 of them had more than one response, leaving only 6 responses to make up the tail.

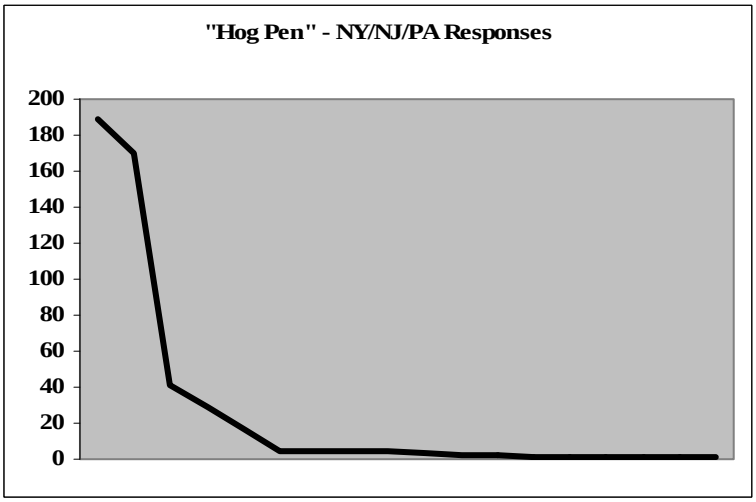


Figure 4: "Hog Pen" NY/NJ/PA Responses

[4] The NY/NJ/PA "hog pen" graph illustrates Distribution Type C, which will be explained further in the "cloudburst" section.

“Meadow” Data

The data for the target lexical item “meadow” also falls into the Distribution Type A category: there were a total of 127 different items and 1,386 total responses, more than half of which were “meadow” and 86 of which were unique items.

Table 2: Top 10 Responses for “Meadow” – All Responses

meadow	872
swale	128
savanna	56
hayfield	37
meadow land	23
prairie	22
bog	18
pasture	17
savanna land	11
old field	8

“Meadow” is the most common response for every subset, although by an even wider margin than for “hog pen”: the average difference between the first and second most common responses for this lexical item is 271. For all subsets except MD/DC/DE/VA/WV and NC/SC/GA/FL, “swale” occurs in the second most frequent position (“meadowland” and “savanna” are the second most common items for these two geographic subsets, respectively). For this data, the lexical item in the third position was also fairly consistent among subsets, with “savanna” appearing in fifty percent of the subsets; the exceptions were Females and NY/NJ/PA (“hayfield”), MD/DC/DE/VA/WV (“bottom”), and NC/SC/GA/FL (“bog”). “Savanna” did not appear in the top 10 responses for either NY/NJ/PA or MD/DC/DE/VA/WV, while the MD/DC/DE/VA/WV and NC/SC/GA/FL subsets also did not have “hayfield” in their top 10 lists, though this lexical item was the fourth most common response for three other subsets. For this data, the level of consistency between subsets dropped below forty percent beginning with

the fourth position and continued lowering with each subsequent position. As with the data for "hog pen," there are eight responses which appear in only one of the eight subsets and these include: "beaver meadow," "hay lot," and "pond" (NY/NJ/PA), "branch," "buffalo wallow," "glade," and "grass bottoms" (MD/DC/DE/VA/WV), and "pasture land" (NC/SC/GA/FL). In this case, the majority of these outlying responses come from the NY/NJ/PA respondents or those from the MD/DC/DE/VA/WV area.

The majority of the subsets had the same items in similar orders for their top 10 lists, but the three geographic subsets showed some differences. For most of the subsets, "savanna" was the third most common response, yet this item did not occur in the top 10 list for either NY/NJ/PA or MD/DC/DE/VA/WV. Similarly, while "hayfield" appeared as the fourth most common response for several subsets, it did not occur at all for MD/DC/DE/VA/WV or NC/SC/GA/FL. Also, of the seven unique responses, four of them come from the MD/DC/DE/VA/WV subset.

The graph for All Responses exhibits an A-curve with no major bumps or dips, which is echoed in most of the other subset graphs, the exceptions being those for Female, NC/SC/GA/FL, and MD/DC/DE/VA/WV Responses. The Female and NC/SC/GA/FL graphs each have identical bumps in the same location, though they are similar to the other subset graphs in all other respects. The MD/DC/DE/VA/WV graph exhibits a more sharply-angled form; however, it still retains a basic A-curve shape. For this subset, there are only 13 unique items (8 of which are single-response items), and the difference between the first and second most frequent responses is 314, which explains the drastic initial drop and the fact that the tail accounts for the majority of the graph's structure. All of these graphs are A-curves with Distribution Type A.

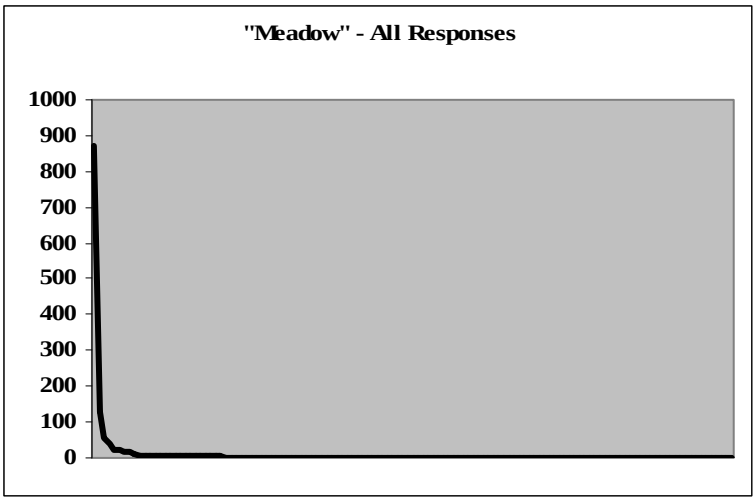


Figure 5: "Meadow" All Responses

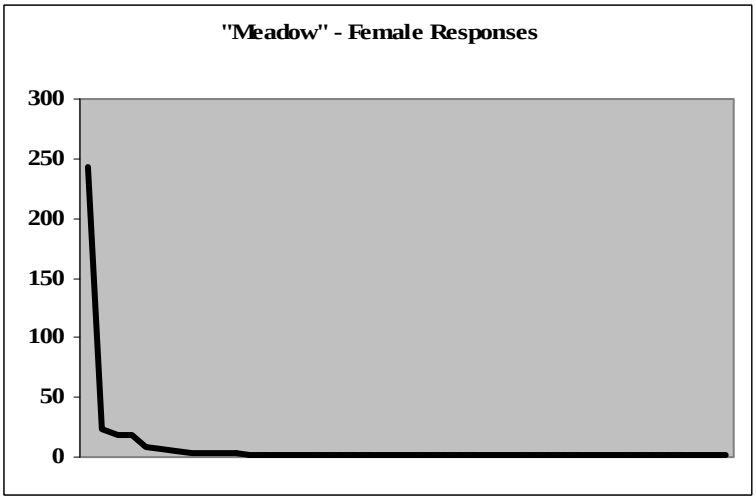


Figure 6: "Meadow" Female Responses

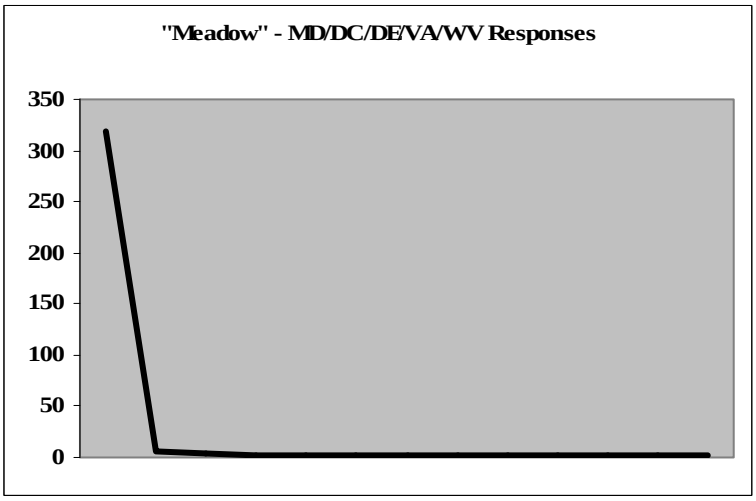


Figure 7: "Meadow" MD/DC/DE/VA/WV Responses

“Swamp” Data

In the “swamp” data set, there were 2,005 total responses and 169 different items, 109 of which were single-response items. This data set also exhibits Distribution Type A, with “swamp” being the most frequent item by a margin of 1,053 responses, “marsh” and “bog” having similar frequencies, and many items with only one or two responses.

Table 3: Top 10 Responses for “Swamp” – All Responses

swamp	1186
marsh	133
bog	103
swamp land	51
pond	37
slough	25
boggy	24
bay	23
branch	18
river swamp	16

As with the two lexical items examined above, all subsets share the same most frequent item, in this case “swamp.” For all but two subsets, “marsh” was the second most common response, the exceptions being NY/NJ/PA and NC/SC/GA/FL, which each had “bog” in the second position; however, these two subsets had “marsh” in the third position, while this position in all the other subsets was filled by “bog”. This level of consistency in the top three responses was not present in the data for either “hog pen” or “meadow.” This consistency begins to lessen with the fourth position and beyond: “swamp land” is fourth most common response for four out of the eight subsets, while for the other subsets, the fourth most common responses are almost all different: “branch” (Age Group 1 and MD/DC/DE/VA/WV), “pond” (Age Group 2), and “fly” (NY/NJ/PA). The trend continues in the fifth position, for which there are six unique responses between the eight subsets: “pond” (Males and NC/SC/GA/FL), “slough” (Age Group 1), “swamp

land" (Age Group 2 and NY/NJ/PA), "bay" (Age Group 3), "boggy" (Females), and "lowgrounds" (DC/DE/MD/VA/WV).

There are fewer outlying responses for this set of data than in the two sets examined above. Variations on "bay" and "boggy" make up most of the lower half of the lists for each subset, and there are only two responses which appear in only one subset each: "frog marsh" (DC/DE/MD/VA/WV) and "hammock" (NC/SC/GA/FL). Again, these outlying responses are present in two of the three geographic subsets, a trend which is found in the data for "hog pen" and "meadow" as well. Also, it is interesting that NY/NJ/PA and NC/SC/GA/FL were the only two subsets with a different frequency order for their second and third most common items, especially considering these two geographic areas are separated by the MD/DC/DE/VA/WV subset, which conformed to the pattern shown by all other subsets.

The All Responses graph for "swamp" also exhibits an A-curve with only a very slight bump after the initial drop; this basic form is echoed in all of the subset graphs for this lexical item, although sometimes the bump is a little less smooth, as in the graph for the MD/DC/DE/VA/WV subset. For this latter set of data, there were only two different lexical items with more than 10 responses, resulting in a much sharper initial drop.

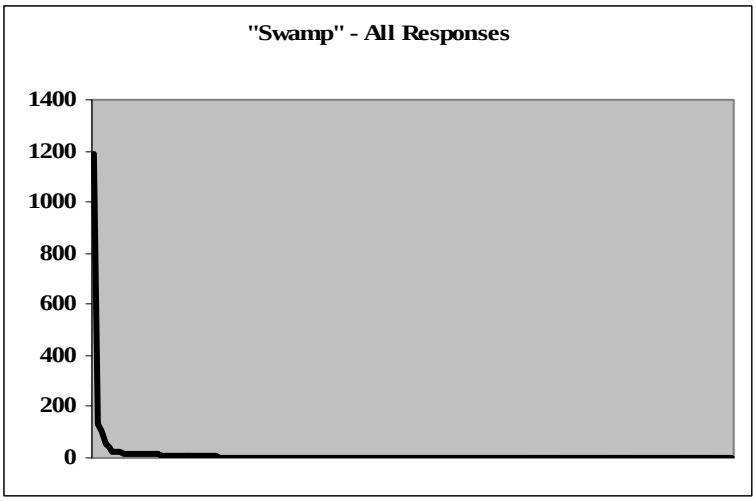


Figure 8: "Swamp" All Responses

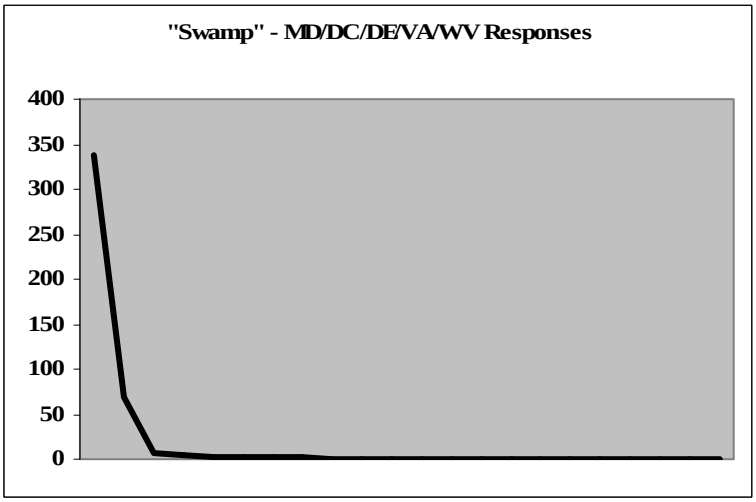


Figure 9: "Swamp" MD/DC/DE/VA/WV Responses

Within the Food domain, the three target lexical items examined were "cobbler," "corn bread," and "pancakes."

"Cobbler" Data

There were 265 different items (159 were single-response items) and 1,710 total responses in the data for "cobbler". This set of data exhibits a different pattern of distribution than the previous data sets, and it also includes a greater number of different items and unique

responses than in any of the other data sets (both preceding and following). The majority of responses were not concentrated on a single lexical item, but were instead divided evenly among the first few items and the frequency decreases gradually, as seen in the top 10 responses chart below.

Table 4: Top 10 Responses for "Cobbler" – All Responses

apple cobbler	130
family pie	115
cobbler	108
apple pie	97
apple dumpling	84
apple pot pie	72
pot pie	65
apple pudding	53
deep dish apple pie	52
peach pie	52

The data for this lexical item shows much less uniformity than that for the items examined above. There are four different most frequent responses among the subsets: "apple cobbler" (Age Group 1, Males, and MD/DC/DE/VA/WV), "apple pie" (Age Group 2, Age Group 3, and NC/SC/GA/FL), "family pie" (Females), and "apple dumpling" (NY/NJ/PA). This variety in responses also occurs for those items in the second position. While "family pie" occurred for both Age Group 1 and NC/SC/GA/FL, and "cobbler" for Age Group 2, Males, and MD/DC/DE/VA/WV), the remaining three subsets each had a different response in this position: "apple dumpling" for Age Group 3, "apple cobbler" for Females, and "apple pot pie" for NY/NJ/PA. Fifty percent of the subsets had either "cobbler" or "apple cobbler" in the third most common position, but again, the remaining subsets each had a different item: "apple pie" (Males), "deep dish apple pie" (NY/NJ/PA), "family pie" (DC/DE/MD/VA/WV), and "peach pie" (NC/SC/GA/FL). While the fourth position for most subsets was filled by a lexical item that appears elsewhere in another subset, the NY/NJ/PA subset had "apple grunt" as the fourth

most common response, an item which does not appear anywhere else in the "cobbler" data. In the fifth position or lower, there are six other items which appear in only one subset: "deep apple pie[5]" (Females), "birds nest" and "crows nest" (NY/NJ/PA), "big pie" and "cut and come" (MD/DC/DE/VA/WV), and "potato pie" (NC/SC/GA/FL). Again, most of these outlying responses are part of the geographic subsets.

Among the most frequent responses, the Female subset was the most different from the other responses, with no mention of the word apple. The youngest respondents and middle states preferred types of "cobbler", while the older respondents and southern states preferred to use the term "pie" in their responses. Although the level of agreement between the subsets was low compared to the previous lexical items examined, the NY/NJ/PA subset showed even less agreement than the other subsets within this data set. For the first through third frequency slots, the NY/NJ/PA subset had an item that no other subset shared in these slots, though the responses did occur lower in other lists. The fourth most frequent item, though, ("apple grunt") did not occur in the top 10 lists for any of the other subsets, and the same is true for the fifth and seventh most frequent items for the NY/NJ/PA subset.

Despite the amount of variation in the responses for this data set, the All Responses graph for "cobbler" still illustrates an A-curve, although it is quite bumpy. This type of A-curve, with a slower decrease than in previous graphs, results from having more than one highly-occurring response (in this instance, there are only 22 responses different between the first and third most common lexical item, and there are 31 items with more than 10 responses each). This type of A-curve illustrates what I have termed Distribution Type B, characterized by a more gradual initial decrease in response frequency and a tail made up of items with only one or two responses each.

[5] There are several occurrences of "deep *dish* apple pie", but only one occurrence of this particular item; the remaining outlying responses are significantly different from the other responses in this data set.

The bumps in the “cobbler” graph occur because pairs of lexical items with the same or similar number of responses are dispersed throughout the data and the tail begins later along the curve.

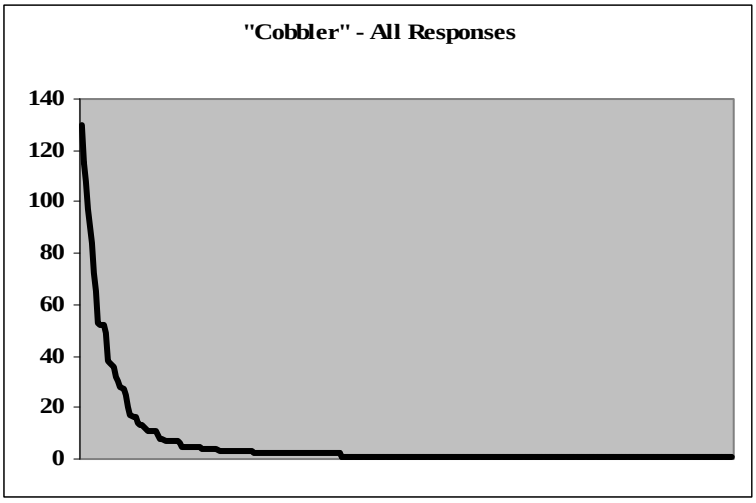


Figure 10: “Cobbler” All Responses

For this data set, all of the subset graphs are practically identical to that of All Responses, as illustrated below with the Age Group 1 graph.

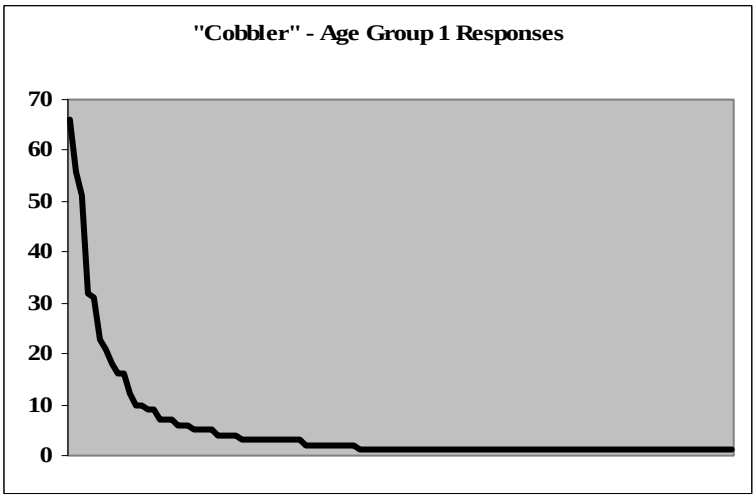


Figure 11: “Cobbler” Age Group 1 Responses

Even the graph for the NY/NJ/PA subset, which showed variation in the frequency charts, has an A-curve of Distribution Type B.

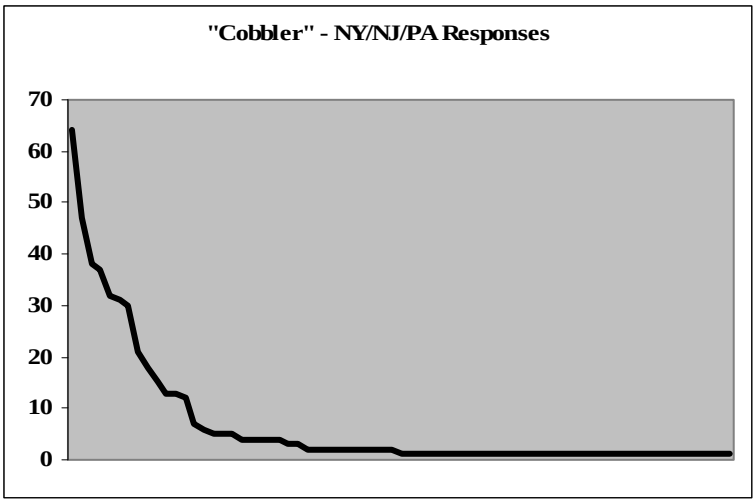


Figure 12: "Cobbler" NY/NJ/PA Responses

"Corn Bread" Data

There were 1,853 responses given for the target lexical item "corn bread," of which 128 were different items. Unique responses account for almost half (81) of the different items. The distribution pattern is that of Distribution Type A: there is one highly-frequent item (although it is not always the most frequent within the subsets), followed by a few other items with high frequencies, followed by a rapid decrease in frequency.

Table 5: Top 10 Responses for "Corn Bread" – All Responses

corn bread	743
corn pone	277
johnny cake	198
Pone	122
pone of corn bread	56
pone bread	48
pone of bread	47
hoecake	34
bread	32
corn cake	32

For this lexical item data, the most common response for every subset except one is "corn bread". For the NY/NJ/PA subset, "johnny cake" is the number one response, although it is

followed closely by “cornbread.” Similarly, the number two most common response for all subsets except NY/NJ/PA is “corn pone” (and this item occurs in the third position for NY/NJ/PA). “Johnny cake” occurs as the third most frequent item for five of the eight subsets, the exceptions being NY/NJ/PA as mentioned above, MD/DC/DE/VA/WV (“pone of corn bread”), and NC/SC/GA/FL (“pone”). All but these last two subsets alternate the same items in the first three slots (“corn bread, johnny cake, corn pone”), but even “pone of corn bread” and “pone” are comparable to these most common responses. For the fourth most frequent item, all of the subsets use a variation of “pone” or “bread,” such as “pone of bread” (Females) and “pone bread” (NC/SC/GA/FL). The same is true for the fifth position, except for that of the NY/NJ/PA subset, in which “corn cake” appears, but variations of “cake” appear lower in the list for the other subsets. Most of the rest of the responses are similar, usually containing some combination of “pone,” “corn,” “bread,” and/or “cake”—for example, “corn dodger” and “hoecake” were fairly common among the subsets. Five items occur only once in the entire data set and include: “corn meal bread,” “Indian bread,” “corn meal muffins,” and “brown bread” (NY/NJ/PA), and “batter bread” (MD/DC/DE/VA/WV). Once again, the outlying responses appear within the geographic subsets, and in this instance, they are made up of mostly NY/NJ/PA responses.

It is interesting to note that although “johnny cake” is only the third most frequent item in the All Responses set, 189 of its 198 occurrences come from the NY/NJ/PA subset, for which it is the most frequent response. This subset also has the two unique items which are the most dissimilar from responses given by any other subset: “Indian bread” and “brown bread”.

The All Responses graph for this set of data is very similar to that of “meadow”: an A-curve with a steep initial drop and then a slight bump.

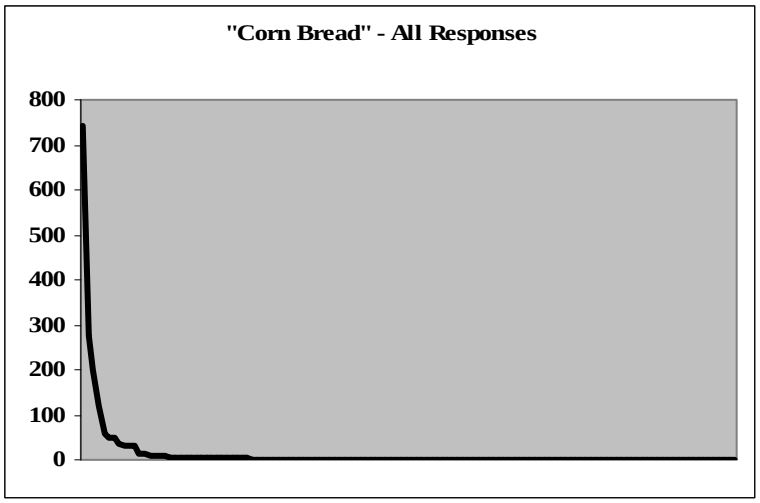


Figure 13: "Corn Bread" All Responses

Most of the subset graphs are very similar, although differences occur in the graphs for Age Group 1 and NY/NJ/PA. The Age Group 1 graph has a smooth curve rather than a bump, as shown below. The responses for this subset seem to drop steadily, rather than rapidly, as in the other graphs.

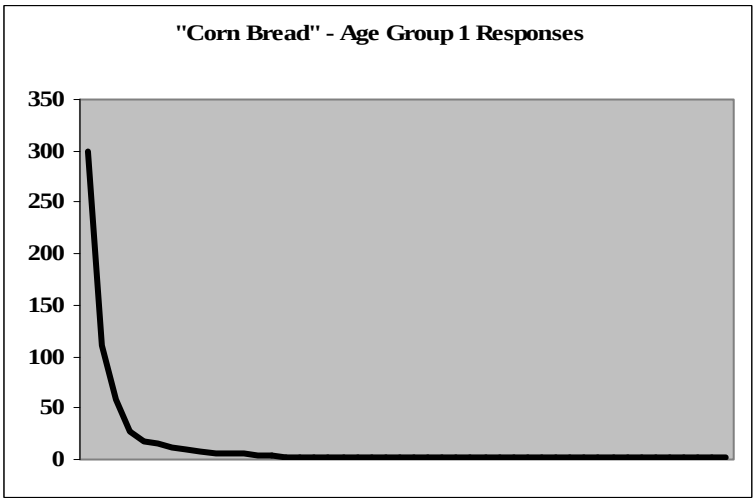


Figure 14: "Corn Bread" Age Group 1 Responses

The graph of NY/NJ/PA responses exhibits a small initial drop (there are only 9 responses different between the first and second most common responses), although after this difference it matches that of the other subsets, with small bumps occurring periodically[6].

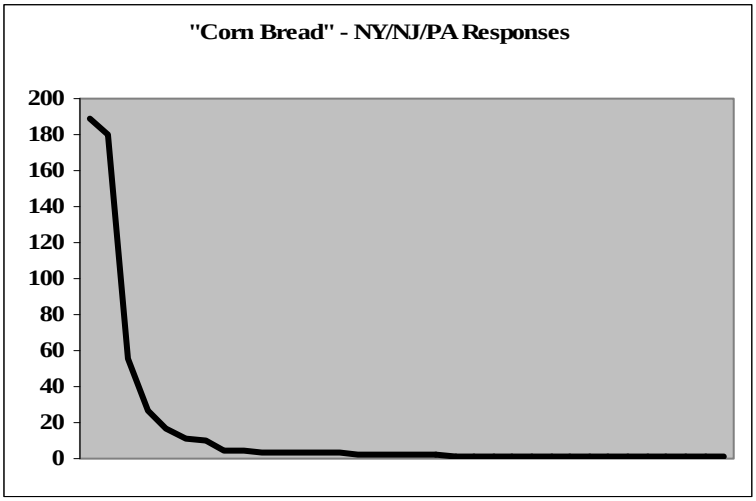


Figure 15: "Corn Bread" NY/NJ/PA Responses

"Pancakes" Data

For this set of data, there were significantly more responses provided than any other data sets (2,660 total), though the number of different responses was only 126, of which 78 were unique items. The pattern of distribution is that of Type A, with "pancakes" being the most frequent response among all subsets, but followed by two items which each had over 300 responses, and then several items with around 200 responses, after which the frequency drops rapidly.

[6] The NY/NJ/PA "corn bread" graph illustrates Distribution Type C, which will be explained further in the "cloudburst" section.

Table 6: Top 10 Responses for "Pancakes" – All Responses

pancakes	757
batter cake	358
flitter	304
fritter	260
flapjack	192
flannel cake	180
griddle cake	149
hotcake	76
slapjack	49
wheat cakes	37

"Batter cake(s)" occurs for five of the eight subsets in the second most frequent position, while "flitters" (Males and MD/DC/DE/VA/WV) and "griddle cakes" (NY/NJ/PA) are the remaining responses. "Flitters" and "fritters" appear in similar positions (between second and fifth) for most subsets, although the NY/NJ/PA subset has "fritters" and "flitters" in the sixth and seventh positions, respectively, while "flannel cakes" and "flapjacks" are in positions three and four. For the other subsets, though, "flannel cakes" and "flapjacks" occur between positions four and seven, except in the NC/SC/GA/FL subset, which has "flannel cakes" in the ninth position. The remainder of the responses are either "cakes" or "jacks" of some sort, such as "wheat cakes, griddle cakes, hot cakes" or "slapjacks, flapjacks". The only outlying response that occurs in this data set is "flitter cakes," which does follow the "cake" pattern just mentioned, but it only occurs in the MD/DC/DE/VA/WV subset. For the most part, although the responses occur in different positions among the various subsets, this set of data is quite uniform and consistent. As the examples above illustrate, even the exceptions follow the basic pattern, and there are no drastically different responses present.

The All Responses graph for "pancakes" is very similar to that of "cobbler," with a series of small bumps occurring throughout the graph, although there are fewer bumps on the

“pancakes” graph. The steep initial drop and the many bumps in the graphs are due to the presence of one highly occurring lexical item, followed by several items with a similar number of responses.

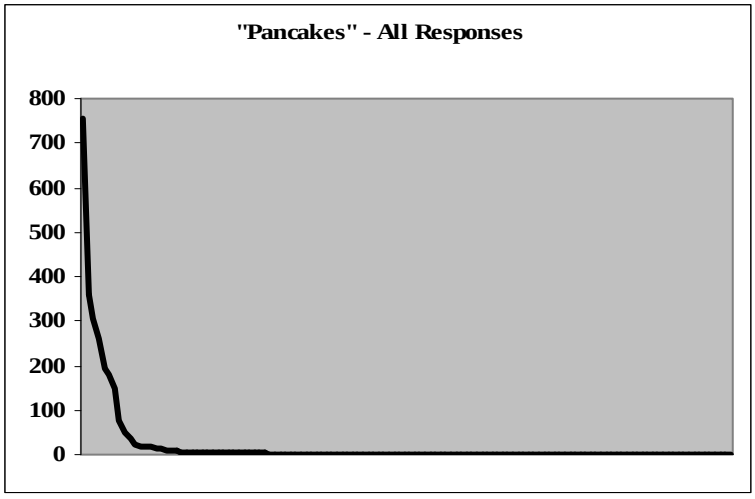


Figure 16: “Pancakes” All Responses

All of the subset graphs conform to this basic pattern, although each graph shows varying numbers of steep and smooth bumps. The graph of Female responses, for example, shows only one significant bump, while that of MD/DC/DE/VA/WV has at least three, as illustrated below.

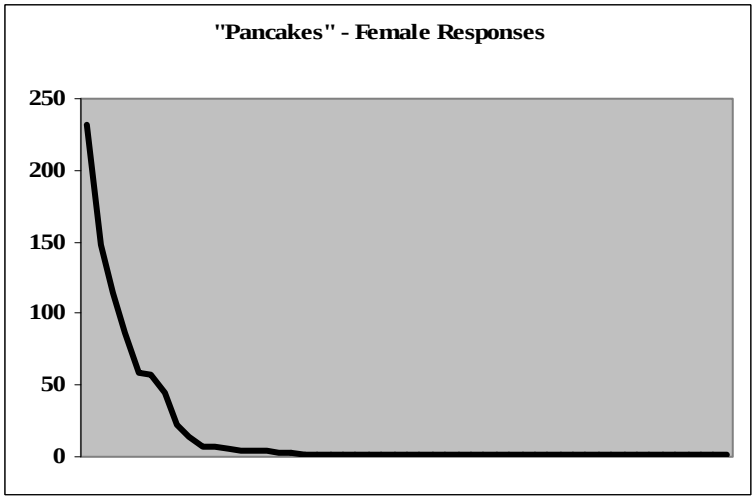


Figure 17: “Pancakes” Female Responses

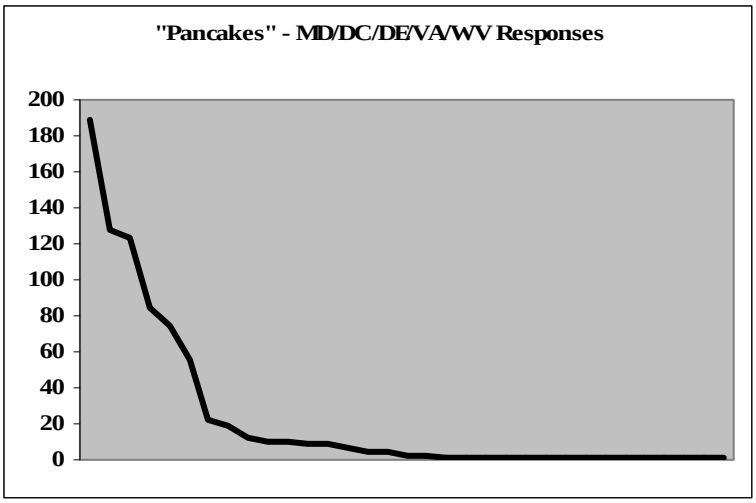


Figure 18: "Pancakes" MD/DC/DE/VA/WV

Within the Home/Household Items domain, the three target lexical items examined were "andirons," "hearth," and "pallet."

"Andirons" Data

For the target lexical item "andirons," there are a total of 1,537 responses and 54 different items, 34 of which were provided by a single respondent. This set of data exhibits Distribution Type A, in that "andirons" is the most frequent item by a margin of 204 responses, but it is followed by three items with over 100 responses each, after which the frequency drops.

Table 7: Top 10 Responses for "Andirons" – All Responses

andirons	553
firedogs	349
dog irons	258
handirons	116
dogs	73
fire irons	66
grate	18
irons	17
fire rocks	16
fender	5

Each of the top 10 lists for this data ends with a lexical item that has five or fewer responses, and there are typically only two or three items with more than 100 responses. "Andirons" is the most common response for every subset except NC/SC/GA/FL, which has "firedogs" in the first position instead. However, "andirons" is in second place for NC/SC/GA/FL, while "firedogs" is the second most common item for all subsets except Age Group 1 and MD/DC/DE/VA/WV ("dog irons"), and NY/NJ/PA ("handirons"). There are only two different items ("dog irons" and "firedogs") in the third most frequent position, and five out of eight subsets shared "handirons" in the fourth most frequent position, the exceptions being "dogs" for Age Group 1 and NC/SC/GA/FL, and "firedogs" for NY/NJ/PA. Again, the fifth most common position is occupied by one of two responses for each subset, either "dogs" or "fire irons." These same two items also appear in the sixth position for six of the eight subsets, while Age Group 1, NY/NJ/PA, and NC/SC/GA/FL have "handirons," "irons," and "grate," respectively, in this position. For the lower positions in each subset's list, "grate," "irons," and "fire rocks" appear most often, but there are four lexical items that only appear in one subset each. These items are: "fire logs" (Females), "sadirons" and "fire basket" (NY/NJ/PA), and "log irons" (MD/DC/DE/VA/WV). As seen in much of the data examined above, NY/NJ/PA and MD/DC/DE/VA/WV have the most responses which only appear once in the entire data set.

Overall, the data for "andirons" is quite similar to that of "pancakes", in that much of the responses are uniform across subsets. However, the exceptions are not as closely related as they were for the "pancakes" data.

The All Responses graph for "andirons" is very similar to that of "pancakes" above, with a steep initial drop and a series of bumps, although this graph has one slightly more exaggerated bump before the tail of the graph begins.

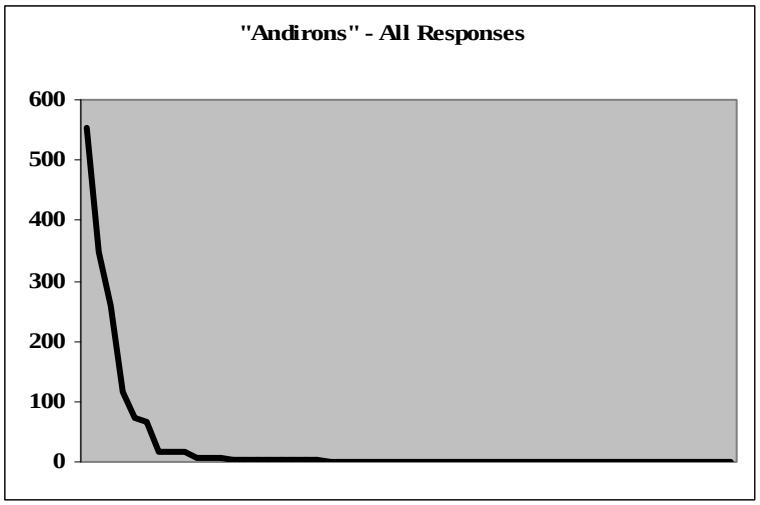


Figure 19: "Andirons" All Responses

Although all of the graphs for this lexical item exhibit A-curves, half of the subset graphs do not quite match the "andirons" All Responses graph. The graph for Age Group 1 has very sharp angles, while the graphs for Age Group 2, Age Group 3, MD/DC/DE/VA/WV, and NC/SC/GA/FL all exhibit Distribution Type B with a gradual initial drop and a long tail; three of these graphs are included below. In the Age Group 3 data, there is a difference of only 112 responses between the first and fifth most common responses; the difference between the first and fifth most common responses for Age Group 3 in the "pancakes" data, in contrast, is 214. This pattern is echoed in the graph of MD/DC/DE/VA/WV responses, for which there is a difference of 109 responses between the first and seventh most common responses.

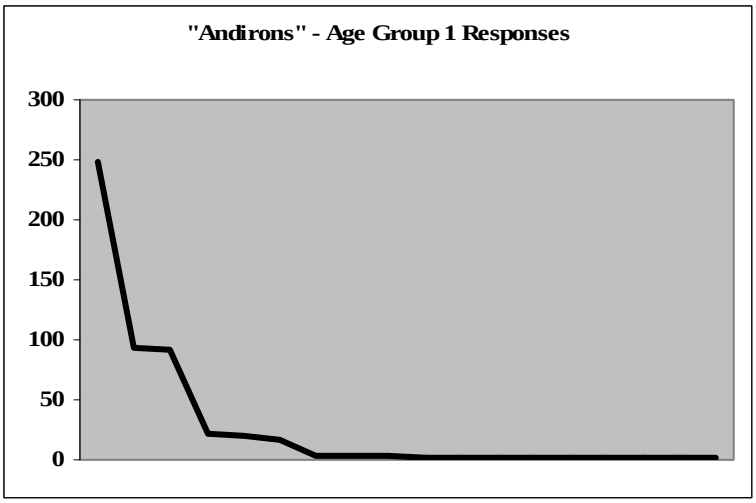


Figure 20: "Andirons" Age Group 1 Responses

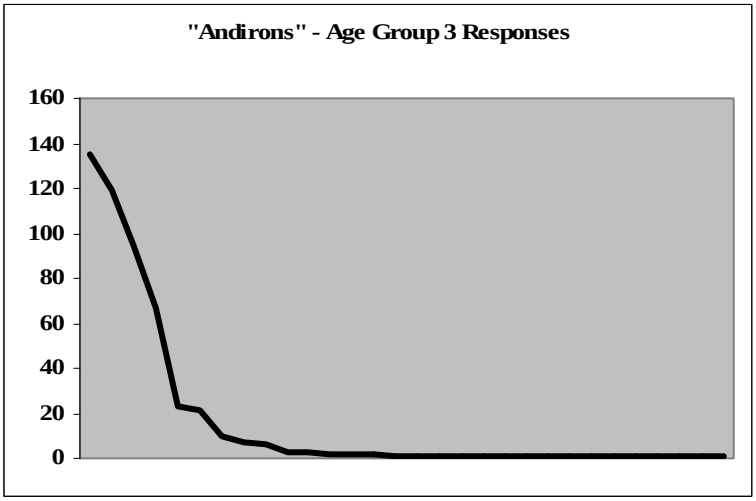


Figure 21: "Andirons" Age Group 3 Responses

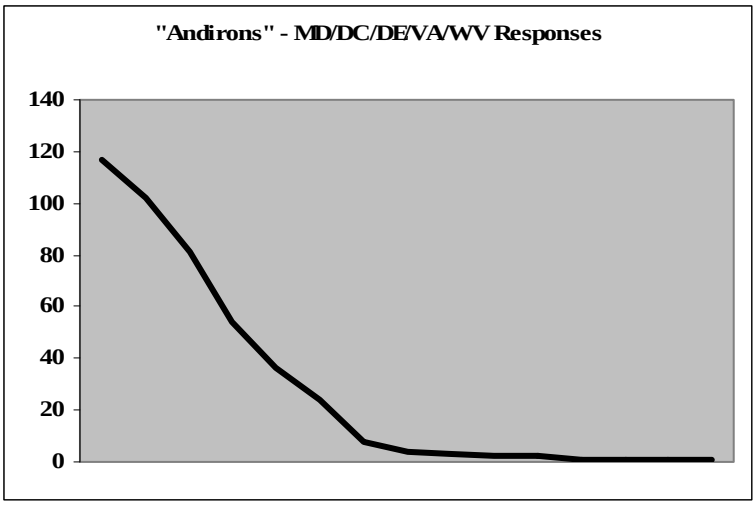


Figure 22: "Andirons" MD/DC/DE/VA/WV Responses

"Hearth" Data

In this set of data, there were 48 different responses within the 1,719 total responses and the vast majority of these were unique items (36 out of the 48 responses). The data for "hearth" exhibits Distribution Type A, with a single highly-occurring response and many less frequent responses, as seen in the top 10 responses chart below.

Table 8: Top 10 Responses for "Hearth" – All Responses

hearth	1430
fireplace	202
fire hearth	20
grate	6
hearthrock	5
chimney hearth	4
stove hearth	4
fires	3
hearth stone	3
ash pan	2

Many of the responses which occur lower in the subset lists for this data were given by only one or two respondents[7]. "Hearth" was the most common response for every subset, typically by a very large margin (the largest difference between the first and second most

[7] Due to the fact that this data set includes so many responses given by a single person, any conclusions drawn from such responses cannot necessarily be generalized across the entire sub-population in which the response occurs.

frequent responses was 1,228 and the smallest difference was 342). "Fireplace" was the second most common response for every subset except MD/DC/DE/VA/WV, which had "fire hearth" in its second position; "fire hearth" appeared for all other subsets in third place, while the MD/DC/DE/VA/WV subset had "hearthrock" as the third most common response. The level of consistency between the subsets drops beginning with the fourth position, in which four unique responses occurred: "grate" (Age Group 1, Females, and NC/SC/GA/FL), "hearth rocks" (Age Group 2 and Males), "chimney hearth" (Age Group 3), "stove hearth" (NY/NJ/PA), and "firerock" (MD/DC/DE/VA/WV—this lexical item does not appear in any other subset lists). For the rest of the lower positions, each lexical item only received one to three responses, and thus will not be examined here.

The All Responses graph for "hearth" exhibits a very steep initial drop, a slight bump, and a long tail, which are the three characteristic features of Distribution Type A.

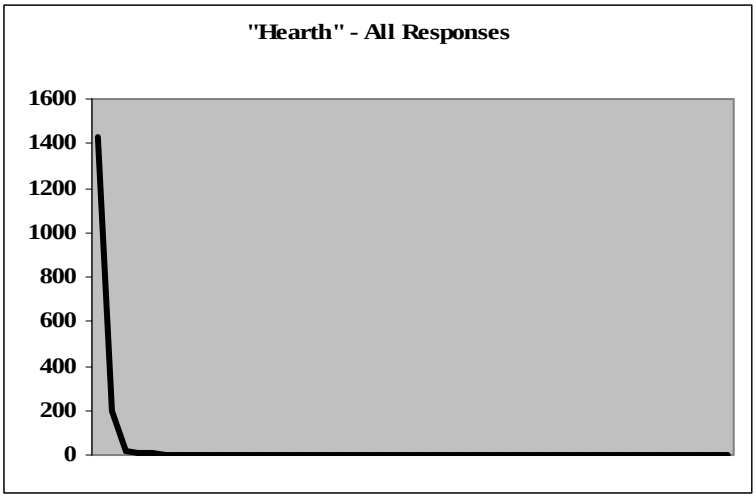


Figure 23: "Hearth" All Responses

The subset graphs for "hearth" all exhibit Distribution Type A, although some of them have much sharper angles instead of smooth curves, as in the graph for Age Group 1. This feature can be explained by the presence of only a few lexical items (typically two or three) with

more than 10 responses each, and between three and eleven items with more than one response each.

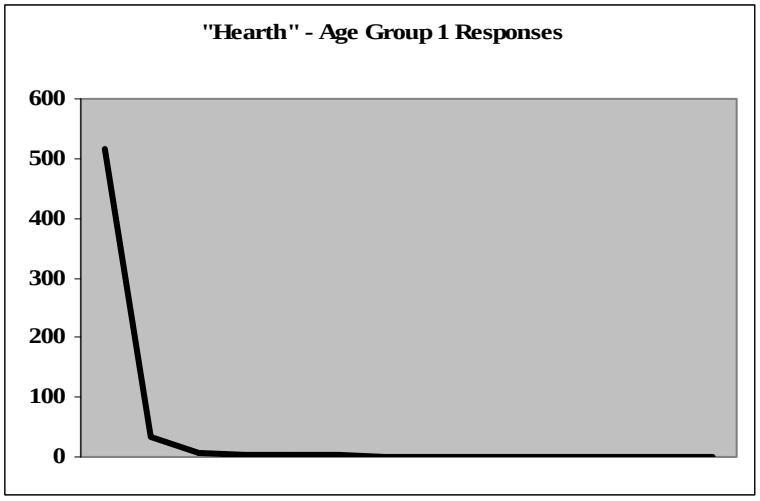


Figure 24: "Hearth" Age Group 1 Responses

"Pallet" Data

For the target lexical item "pallet", there were only 1,009 total responses and 45 different responses. Twenty-six of these responses were unique items. This data again illustrates an A-curve with Distribution Type A.

Table 9: Top 10 Responses for "Pallet" – All Responses

pallet	652
bunk	165
lodging	38
mattress	30
featherbed	24
trundle bed	19
lodge	12
bed on the floor	10
shakedown	6
bed	4

As with the data for "hearth", there were only a few unique lexical items given as responses for this set of data, and many of the lower positions are occupied by items with only

one to three responses each. The most common response for all but one subset is "pallet," for which the NY/NJ/PA subset instead has "bunk" ("pallet" does not occur anywhere in the top 10 list for NY/NJ/PA). "Bunk" is the second most common response for all subsets except NY/NJ/PA ("featherbed") and NC/SC/GA/FL ("lodging"). In the third position, "lodging" appears for three subsets, while these other items appear for the other five subsets: "mattress" for Age Group 3, Males, and NC/SC/GA/FL, "shakedown" for NY/NJ/PA ("lodging" does not appear in the top 10 list of this subset), and "lodge" for DC/DE/MD/VA/WV (again "lodging" is nowhere in the top 10 responses). "Trundle bed" occurs in the fourth most frequent spot for three out of eight of the subsets, while the other five subsets each have unique responses: "mattress" (Age Group 1), "featherbed" (Age Group 3), "lodging" (Males), "straw tick" (NY/NJ/PA), and "spread" (MD/DC/DE/VA/WV). "Bed on the floor" occurs somewhere in all the lists except for the NY/NJ/PA subset ("floor bed" does occur in the tenth spot for this subset, but it is only given by one respondent). Disregarding those items with only one response each, there are five outlying responses which appear in only one subset each[8]: "bollick," "cord bedstead," and "feather tick" (NY/NJ/PA), and "made-down bed" and "on the floor" (NC/SC/GA/FL).

Overall, the NY/NJ/PA subset had very different answers from the rest of the subsets. Of the 131 responses in this subset, 104 were for the most common lexical item, "bunk", and several of the items that occur in the NY/NJ/PA top 10 list do not occur in any of the other sets ("bollick, cord bedstead, feather tick"—each had two respondents in the NY/NJ/PA data). All of the other subsets had "pallet" as the most common response, but this item does not appear at all in the NY/NJ/PA top 10 list.

[8] Each of these five items was provided by only two different respondents, so any conclusions drawn should be carefully backed up with additional evidence.

The All Responses graph for “pallet” looks very similar to that of “hearth,” as an A-curve with a very steep initial drop, a slight bump, and a long tail. Also analogous to the “hearth” data, several of the subset graphs for “pallet” are A-curves with sharper angles, as illustrated by the Female Responses graph. For this subset, there were only three items with more than 10 responses each, and eight others with between two and eight responses; however, each of these graphs exhibits Distribution Type A.

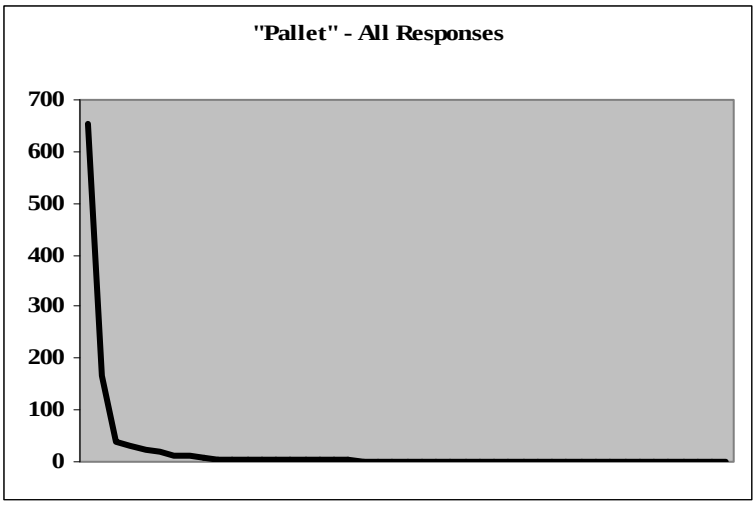


Figure 25: “Pallet” All Responses

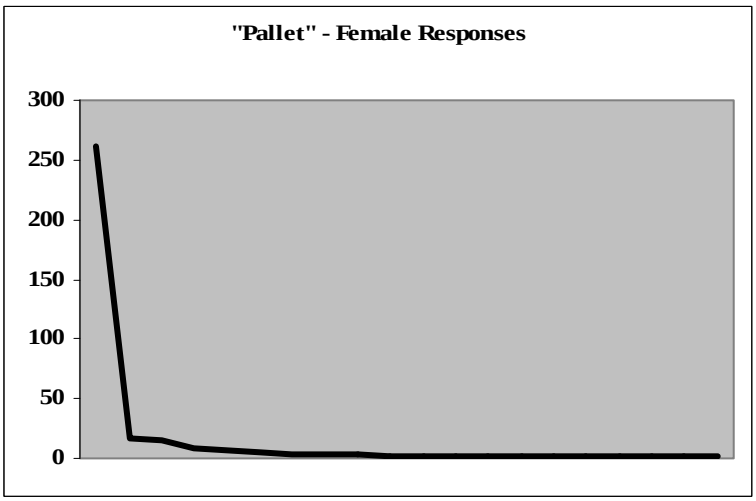


Figure 26: “Pallet” Female Responses

Within the Weather domain, the three target lexical items examined were “cloudburst,” “dry spell,” and “steady drizzle.”

“Cloudburst” Data

In the “cloudburst” data set, there were 1,938 total responses, made up of 187 different items, 114 of which were single-response items. This data exhibits a slight variation on Distribution Type B: there are two highly-frequent items with a similar number of responses, followed by less frequent items, as shown in the frequency chart below. I have labeled this pattern as Distribution Type C, and it is characterized by a very gradual initial drop in frequency, followed by a steep drop, and then the long tail that characterizes all A-curves.

Table 10: Top 10 Responses for “Cloudburst” – All Responses

cloudburst	345
downpour	332
heavy rain	171
hard rain	146
gully washer	119
big rain	84
flood	84
pourdown	61
hard shower	57
heavy shower	47

For this data set, there is not a single most common response; instead, the subsets are divided between two options: “cloudburst” was preferred by Age Group 3, Males, and NY/NJ/PA, while “downpour” was most common for Age Group 1, Age Group 2, Females, MD/DC/DE/VA/WV, and NC/SC/GA/FL. While more than half of the subsets had “downpour” as the most common response, “cloudburst” was the most common item overall, although by only 13 responses. Every subset that had “downpour” as the most common response had “cloudburst” as the second most common response. Of the subsets with “cloudburst” as the most common response only Age Group 3 did not have “downpour” in the second most frequent

position (“heavy rain” instead appeared). For the third most common response, most of the subsets had either “heavy rain” or “hard rain,” with two exceptions: “downpour” (Age Group 3) and “gully washer” (NC/SC/GA/FL). A similar pattern is true for the fourth position, in which “hard rain” or “heavy rain” appeared for all subsets except NY/NJ/PA (“heavy shower”). All subsets except two had “gully washer” for the fifth most frequent response: NY/NJ/PA instead had “hard rain” (“gully washer” does not appear in the top 10 list for this subset at all) and NC/SC/GA/FL had “hard rain” instead. As in previous data sets, the consistency between subsets begins to lessen further down the list, although for this data, it does not happen until the sixth position. Although “big rain” and “flood” were commonly in the sixth and seventh positions, “hard shower” was also common, and “pourdown” appears in the sixth position for MD/DC/DE/VA/WV. “Pourdown” occurs lower in the list for every other subset, in either the eighth or ninth positions. The rest of each list is made up mostly of items mentioned above, in varying positions, although five unique items occurred in only one subset each: “shower” (Females), “pouring rain” (NY/NJ/PA), “pouring down rain” (MD/DC/DE/VA/WV), “trash mover” and “squall”(NC/SC/GA/FL)—of these outlying responses, only the two occurring in the NC/SC/GA/FL subset are significantly different from the more common responses.

Overall, there is a high degree of consistency in this set of data, with only two significantly different responses provided. Although the NC/SC/GA/FL subset had the two most significantly different outlying responses, no single subset seemed to be drastically different from the others.

The All Responses graph for “cloudburst” has a small initial drop, similar to the NY/NJ/PA subset of cornbread, but after the drop it has a series of small bumps, like the graph of

"cobbler." The small initial drop is found in the graphs for Age Groups 1 and 2, while almost all of the graphs show many bumps.

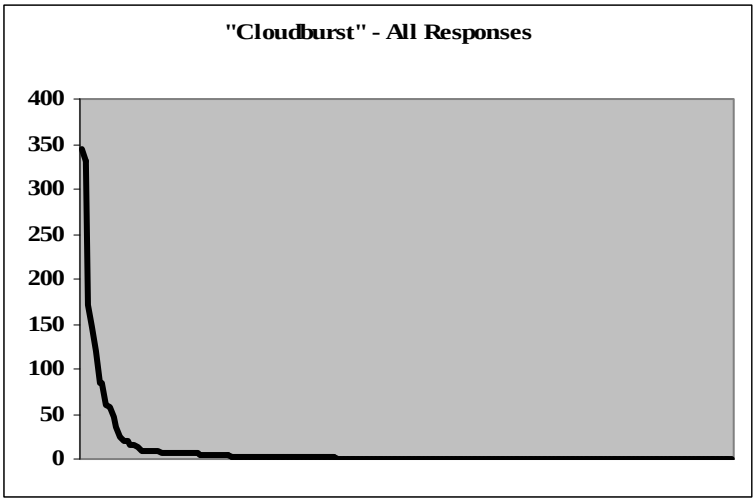


Figure 27: "Cloudburst" All Responses

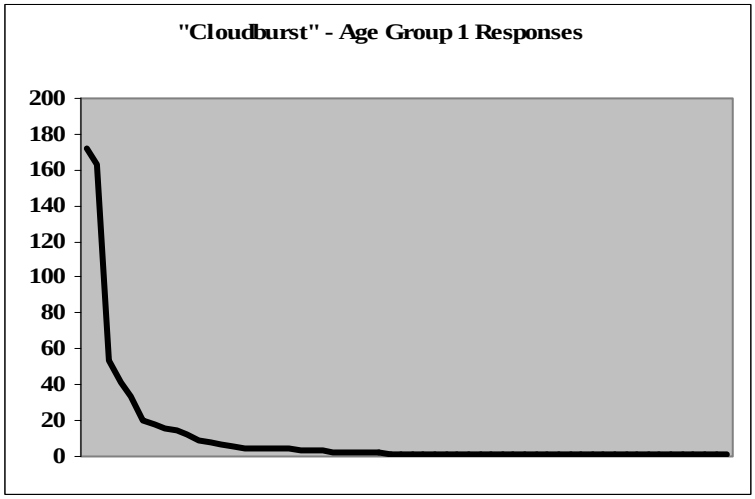


Figure 28: "Cloudburst" Age Group 1 Responses

"Dry Spell" Data

Within this data set, there are 1,953 total responses and 39 different items, 22 of which are unique responses. This data set has the smallest number of different items, as well as unique responses, and this is not due to a smaller number of responses. The "pallet" data set had more than 900 fewer responses than the "dry spell" data does. Although the margin between the first

and second most common responses is a bit larger than in the "cloudburst" data[9], this set is characterized by Distribution Type C: the two most frequent items have very similar frequencies, followed by a rapid decrease in frequency for the following responses.

Table 11: Top 10 Responses for "Dry Spell" – All Responses

dry spell	906
drouth	720
drought	164
dry weather	38
dry time	25
dry drouth	15
dry drought	14
dry season	10
wet spell	9
spell of dry weather	8

For this data set, there is a difference between subsets in how many unique responses were provided. While some subsets have lexical items with only one response in several top 10 positions (Age Group 1, for example), there are other subsets whose tenth most common item had as many as five responses (Males, for example). "Dry spell" is the most common response for all subsets, "drouth[10]" is second most frequent for all subsets, and "drought" is the third most frequent response for all subsets. Additionally, there are only two different lexical items that appear in the fourth position: "dry weather" for Age Group 1, Age Group 2, Males, MD/DC/DE/VA/WV, and NC/SC/GA/FL, and "dry time" for Age Group 3, Females, and NY/NJ/PA. Although there are a handful of different responses in the fifth position, they are all very similar: "dry season" (Age Group 1), "dry drought" (Age Group 2 and NC/SC/GA/FL),

[9] Several "dry spell" subsets show larger margins between the first and second most frequent responses than do the "cloudburst" subsets which were characterized as Distribution Type A. This is because many of the "cloudburst" subsets' most frequent responses were in the low 1- to 200 range, while the "dry spell" data contains items with frequencies in the 5- to 600 range. Therefore, although the actual margin of difference is larger, the relative difference is smaller.

[10] "Drouth" is identified as a secondary spelling of "drought" in the Oxford English Dictionary. The most recent quotation given with the spelling "drouth" is from 1865 ("Drought, Drouth").

"dry weather" (Age Group 3 and Females), "dry time" (Males and MD/DC/DE/VA/WV), and "spell of dry weather" (NY/NJ/PA). The rest of the responses in the top 10 lists are quite similar, using the word "dry" either by itself or in combination with "drouth, time, season, drought," etc. One exception is "wet spell[11]," which is the exact opposite of the target lexical item, but occurred in four different subsets: Age Groups 2 and 3, Males, and NC/SC/GA/FL. The three items which occurred in only one subset each were "a spell of weather," "clear spell" (both from the NY/NJ/PA subset) and "draft" (DC/DE/MD/VA/WV), although these outlying responses were provided by only one respondent each. Overall, this set of data shows a great amount of consistency in responses from the various subsets. The three most common responses were all the same for every subset, and even for the fourth spot, only two different responses were given.

Overall, this set of data shows a great amount of consistency in responses from the various subsets. The three most common responses were all the same for every subset, and only two different responses were given even in the fourth spot.

The graph of All Responses for "dry spell" illustrates an A-curve with Distribution Type C: the graph exhibits a small initial drop followed immediately by a steeper drop, and then one significant bump at the end of the steep drop before the tail of the graph begins. This pattern is echoed in most of the subset graphs, with varying degrees of small and steep drops. For example, the graph for the MD/DC/DE/VA/WV subset has a very small initial drop (only a three-response difference between the first and second most frequent responses) followed by a steeper drop, as seen below.

[11] An argument could be made that "wet spell" should have been labeled an "Inappropriate Response;" however, I made the decision to keep it in my data since it was provided by nine different respondents and occurred in the top 10 lists of several subsets.

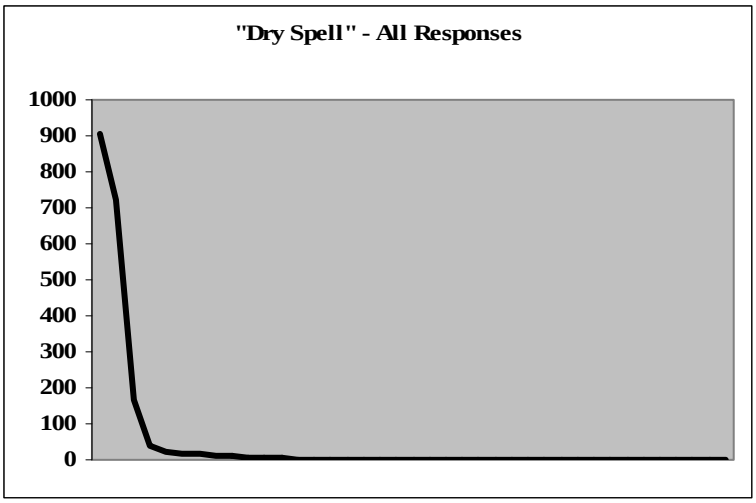


Figure 29: "Dry Spell" All Responses

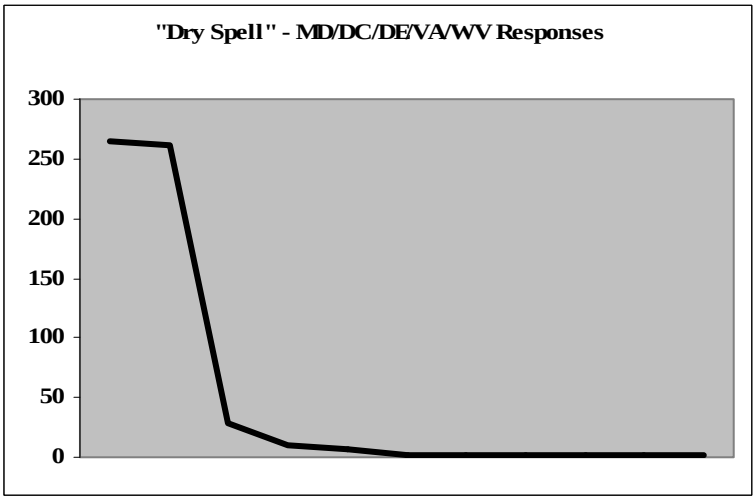


Figure 30: "Dry Spell" MD/DC/DE/VA/WV

The Females graph (shown below) and the NC/SC/GA/FL graph are both missing small initial drops and exhibit A-curves with Distribution Type A instead of Type C.

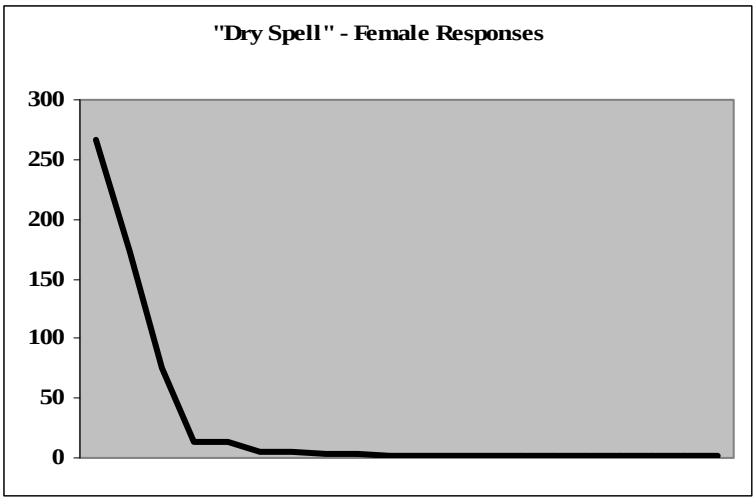


Figure 31: "Dry Spell" Female Responses

"Steady Drizzle" Data

For the target lexical item "steady drizzle", 1,870 responses were provided, 132 of which were different items. Of these responses, 85 were unique items, provided by only one respondent each. This data set exhibits Distribution Type A, with "steady drizzle" being the most frequent response by a large margin.

Table 12: Top 10 Responses for "Steady Drizzle" – All Responses

steady drizzle	796
drizzle	286
shower	233
sprinkle	173
mist	51
light rain	27
steady rain	27
light shower	26
steady	17
misting rain	12

For this data set, only one subset (MD/DC/DE/VA/WV) had fewer than 10 unique responses—all of the other subsets had many more variants in their top 10 lists. "Steady drizzle" is the most common response, typically by a large margin, for every subset except

NC/SC/GA/FL, which has "drizzle" instead. "Steady drizzle" appears in the second most frequent position for NC/SC/GA/FL, while "drizzle" is the second most common response for all other subsets, except NY/NJ/PA ("shower" is second for this subset). "Shower" is the third most common response for all other subsets, except MD/DC/DE/VA/WV, which had "spitting" instead ("shower" does not appear in the list for this subset at all). "Sprinkle" is the fourth most common response for all subsets except MD/DC/DE/VA/WV ("splitting snow" occurs instead, although this item was provided by only one respondent, and "sprinkle" does not appear in the top 10 list for this subset). Similarly, "mist" is the fifth most common response for every subset except MD/DC/DE/VA/WV ("steady grizzle," provided by one respondent). More than half of the subsets[12] have "light rain" in the sixth most frequent position, while the remaining three subsets have either "light shower" (Age Group 3 and Males) or "steady rain" (NY/NJ/PA). These three responses also appear in the seventh and eighth most frequent positions, along with the addition of "drizzling rain" (seventh for Age Group 1), "drizzly" (eighth for Age Group 1), and "steady" (eighth for NC/SC/GA/FL). For the remainder of each list, most of the responses are very similar to those given above, such as "misting rain" or "drizzling rain". There were no responses, other than the two given by one respondent each in the MD/DC/DE/VA/WV subset, which occurred in only one subset.

The MD/DC/DE/VA/WV subset was quite different from the other subsets in several aspects. First of all, there were only five different items in the frequency chart, two of which were provided by only one respondent each[13], and all of the other subsets included more than 10 different responses each. The second most common response overall was "drizzle" and its

[12] The MD/DC/DE/VA/WV subset is not considered in the remaining data analysis for this lexical item because it only had five unique responses in its list.
[13] These two unique responses will not be considered in this section.

assorted variants (“drizzles, drizzly, drizzling”), yet the MD/DC/DE/VA/WV respondents provided “drizzling” as the only variant of this lexical item. Finally, this subset included an item (“spitting”) that did not occur in any of the other subsets and is significantly different from all of the other responses provided.

The All Responses graph for “steady drizzle” has fairly large bump immediately following the initial steep drop, a pattern which is found in most of the subset graphs. The bump appears because of a decline in the rate of frequency change: for the All Responses graph, there is a difference of 53 between the second and third most common responses, and a difference of 60 between the third and fourth most common responses, as compared to the margin of 510 responses different between the first and second most frequent items. All of these graphs illustrate Distribution Type A.

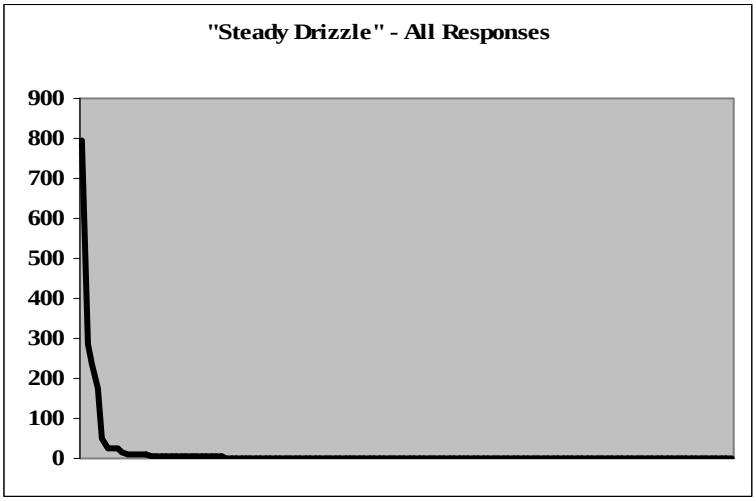


Figure 32: “Steady Drizzle” All Responses

The graph for the MD/DC/DE/VA/WV subset is an A-curve with Distribution Type A, characterized by one highly-frequent item, a steep drop in frequency, and a tail made up of single-response items. However, this subset graph has a truncated tail (made up of only two items) because the subset had only five unique responses given.

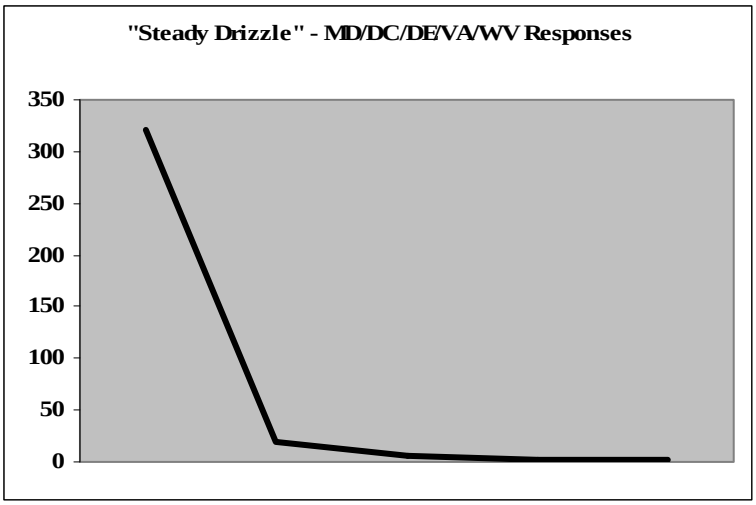


Figure 33: "Steady Drizzle" MD/DC/DE/VA/WV Responses

CHAPTER 5
DISCUSSION

Bybee (Organization, Emergence), Kretzschmar (Dialectology, "Distributional," "Neural", Linguistics), and Cameron and Larsen-Freeman ("Complex," "Research") have been among the most vocal proponents for a new theory of language variation. Additionally, Kretzschmar (Dialectology, "Distributional," "Neural," Linguistics), Burkette, Hoover, and Johnson have all noted that lexical variation data (both from the 1930 LAMSAS surveys and from more recent surveys) exhibits a strikingly regular pattern: when graphed, the data always exhibits an A-curve. The works of these respective authors are valuable contributions to the idea of language as a complex system, although no one to date has analyzed such lexical data from various *sub-samples* of the entire survey population and provided the corresponding frequency graphs. My own examination is largely based off Kretzschmar's 2009 book *The Linguistics of Speech*. With the analysis of 12 different lexical items (and 8 different subsets within each) from different domains of knowledge, this study replicates the studies cited above and provides ample evidence for the idea that language is an example of a complex system and should be analyzed according to a new set of principles. Clearly these lexical items are similar in many ways, yet they also show distinctive differences; both the similarities and the differences must be examined in order to present a clear picture of what is going on in human speech.

The four major assumptions about language which characterize this new approach to the study of language variation are outlined by Kretzschmar in *The Linguistics of Speech* (8) and each will be examined in turn using the LAMSAS data summarized above. Additionally, the

graphs of these lexical items will be examined to determine if they exhibit a "power law" distribution (Kretzschmar, "Distributional" 378) and the characteristics of scaling.

It is a significant observation that language exists as part of a continuum, rather than a series of binary options. For the lexical items examined above, there is an average of 116 different responses per target lexical item[14]. Even for the lexical item with the smallest number of different items provided ("dry spell"), there are still dozens of variants. There is a wide range of choices in the names for even concrete objects, which is clear evidence for the existence of a linguistic continuum. One is not required to use either "apple cobbler" or "family pie"; instead, there are at least 265 other options. The data from the "hog pen" set is particularly indicative of the idea of the linguistic continuum: within the top 10 responses list for every subset, variants including the words "hog" and "pig" can both be found, as well as variations of "pen," "lot," "house," etc. Neither "hog" nor "pig" is entirely dominant within any single subset, nor among the entire survey population.

The idea of the linguistic continuum is also related to the assumption that extensive variation exists in all features of language at all times (Kretzschmar, Linguistics 8). Although the average person recognizes that language is variable, it is unlikely that he could predict the existence of over 200 variants of a single lexical item. As Kretzschmar explains, "It is one of the key findings for the linguistics of speech of large-scale surveys like LAMSAS, that language is ever so much more variable than any individual could predict from personal experience" (Linguistics 93). Even within subsets, such as geographic areas, a large number of variants still exist. For example, there are over 125 different items provided for the target lexical item

[14] There is a wide margin between the lexical items with the most and fewest number of different items and unique responses, though. The "cobbler" data set had the greatest number of different items (265) while there were only 39 different items (the smallest number of all the data sets) provided for "dry spell".

"swamp" within the NC/SC/GA/FL subset alone, and 24 of these items were provided by more than five respondents each.

The third assumption examined here is that regional and/or social proximity is a significant factor to consider when investigating language variation (Kretzschmar, Linguistics 8). While Kurath was generally concerned with drawing distinct language boundaries between "speech areas" (v), such boundaries do not capture the true facts about language variation. Instead, variants of lexical items, as well as other features of language, tend to cluster together, and such bunches of features are difficult to illustrate with lines on a map. The LAMSAS website includes density estimation maps, which illustrate this clustering effect quite clearly by showing only a single item at a time. Four maps from the "andirons" data set are shown below.

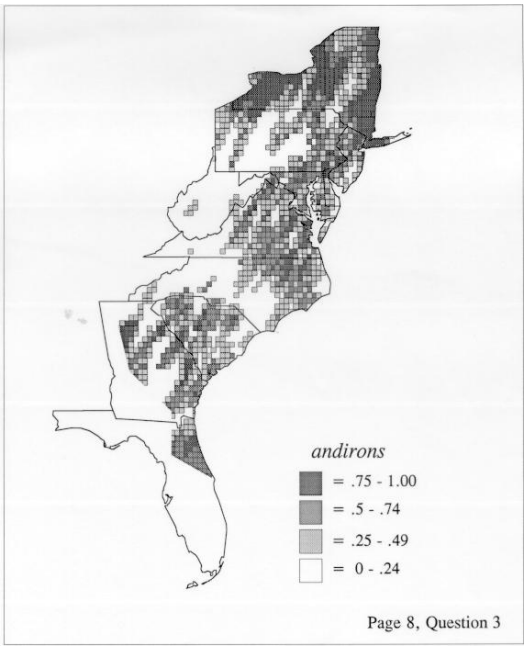


Figure 34: "Andirons" (Density Estimate Map: andirons)

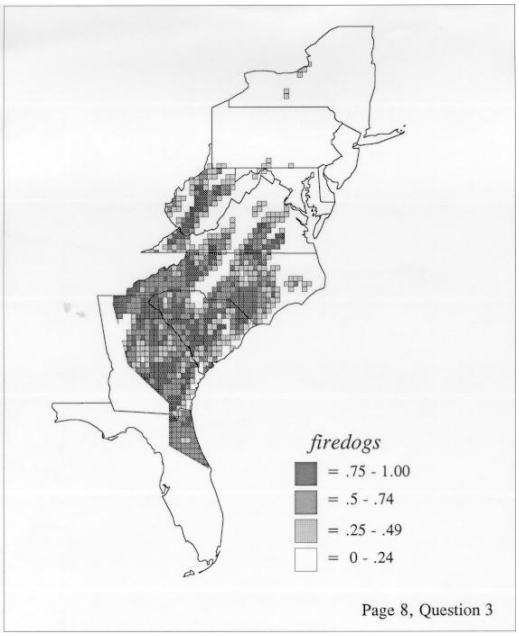


Figure 35: "Firedogs" (Density Estimate Map: firedogs)

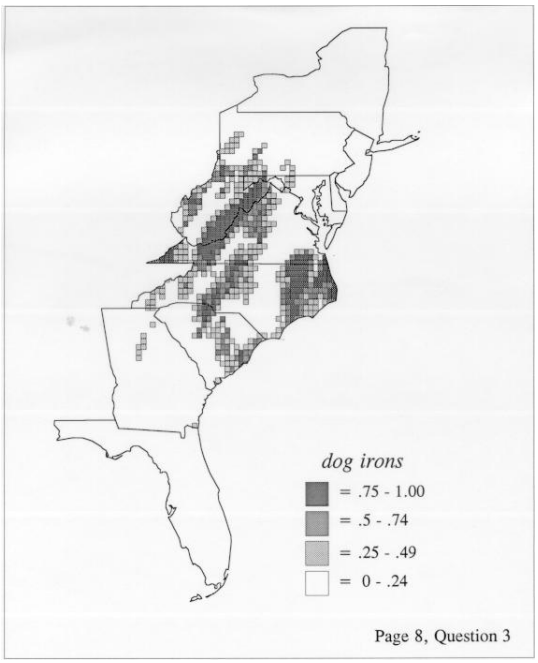


Figure 36: "Dog Irons" (Density Estimate Map: dog irons)

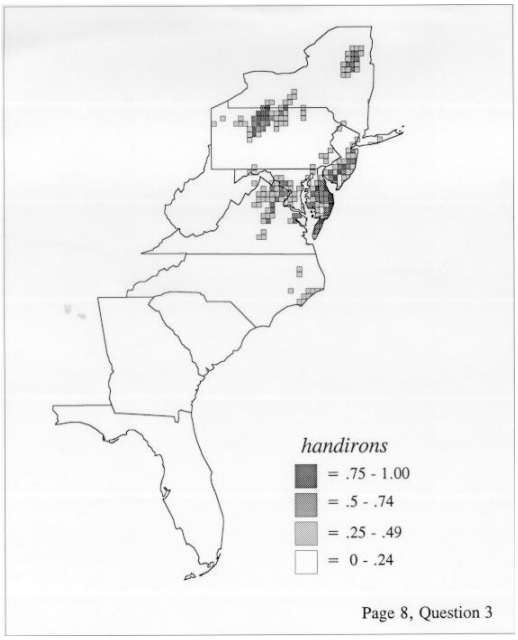


Figure 37: "Handirons" (Density Estimate Map: handirons)

Such evidence based on regional and social proximity "clearly shows that people talking together in localities at least tend to know and to use the same variant for a feature more often than speakers who live far apart" (Kretzschmar, Linguistics 114). At the same time, even within areas where a certain feature is extremely common, there will be pockets in which it is less frequent, again illustrating that language is not a binary phenomenon. However, it is rare that a feature which is extremely common in one subset completely absent in another; instead, "most of these features are at best characterized by changes in local frequency of their variants, not by absolute limits of occurrence" (Kretzschmar, Linguistics 69). For example, in the "pallet" data set, the most common response for all but one subset is "pallet," an item which does not occur anywhere in the top 10 responses list for the NY/NJ/PA subset, but which does occur lower in the list. "Pallet" is the most frequent response by a margin of 369 responses in the NC/SC/GA/FL data and by a margin of 188 responses in the MD/DC/DE/VA/WV subset.

Proximity is also closely related to the fourth assumption, which states that "differential frequency…[is] a key factor in linguistic production both in regional/social groups and in collections of text corpora" (Kretzschmar, Linguistics 8). It is important to take note of proximity-based similarities without falling into the trap of assuming that everyone within a certain geographic or social group will use the same linguistic features. As Kretzschmar explains:

> The perception of aggregated collections of "normal" variants for many features at some level of scale, whether it is a geographical region, a social group, or a text type, creates what we (after Günther et al. 1996, a group of physicists) can call "observational artifacts". That is, dialects as linguistic systems are objects whose existence comes from our perception of reality, not from reality itself… (Kretzschmar, "Dialectology" 20).

In a related paper, Kretzschmar goes on to suggest that the A-curve plays an important function for speakers of a language, even if only subconsciously:

> [S]peakers use the A-curve in order to perceive what is "normal" or "different" for regional or social groups and for text types: the most frequent variants are perceived as "normal" or "expected," and less frequent variants are perceived as "different." What matters for perception is the great difference in frequency between top-ranked variants on the A-curve vs. the lower ranked variants… Moreover, since particular variants are more or less frequent among different groups of people or types of discourse, variants come to be associated with the speaker groups or discourse types by means of these perceptions. ("Neural" 339)

These “observational artifacts” may provide a clue as to where certain popular ideas about language originate. As can be seen in all of the frequency graphs, the distribution of lexical items always exhibits an A-curve—certain items are more frequent than others, but no subset ever includes only one or two items. The MD/DC/DE/VA/WV subset in the “steady drizzle” data set comes closest, but even within this subset there are three different items with more than five responses each, and the graph is still an A-curve.

The notion of a “dialect” is especially problematic according to the linguistics of speech. Dialects “are not ‘natural,’ well-bounded rule systems in the reality of speech production… [A]ny dialect we name actually exists as an observational artifact that comes from our perceptions of the available variants,” and generalizations “always misrepresent…the actual distribution of variants in the group” (Kretzschmar, “Dialectology” 20). Additionally, the most frequent features within a certain subset may be labeled as “standard” or “normal” for that subset, while less commonly used features may be “relegate[d]…to identity marking” and become tied to certain “social or regional identities,” (Kretzschmar, “Distributional” 398-399). Contrary to popular belief, even the most frequent lexical items are not necessarily free from social or regional connections, as Kretzschmar explains:

> [T]here is no particular reason to privilege the most common variants as ‘systematic’ or to relegate less-common variants to identity marking; all of the variants on the A-curve are actually just as relevant for inclusion in the system… It is highly likely that any variant, whether common or rare, will at some subsequent time change in frequency, that some new variant will rise to the top of the curve for a feature…, and that new social and regional associations with variants will form. What is truly stable and systematic about this situation is the

> curve itself, not any perceived system of arrangement of variants" ("Distributional" 399).

Observational artifacts and the perception of dialects are related to saliency, the idea that some features stand out and are more prominent and noticeable than other features. Although it is not widely known that linguistic features exhibit A-curves when graphed, most people do in fact perceive "ranked frequencies in speech according to the A-curve, at whatever scale in any dimension, allow[ing] speakers to identify a feature variant at or near the top of the A-curve for some category as 'right' for that category" (Kretzschmar, Linguistics 206). Observational artifacts emerge when a certain frequent variant becomes associated with a particular social or regional group and comes to define that group, such as the idea that all American Southerners must use "y'all" for "you plural" (Kretzschmar, "Neural" 341). It is common for speakers of a language to use the most salient features of a particular group of speakers to characterize that group, as Kretzschmar explains:

> When we notice that one variant appears to be the most frequent one for a group of speakers, we commonly make the generalization that the variant is "normal" for everybody in that group, and that everybody habitually uses it. Taking many features together at once, we make the further generalization that a dialect exists, composed of the normal use of some set of particular features by its speakers… This kind of generalization is an observational artifact because it always misrepresents the actual distribution of variants in the group—which in fact, as we now know, always has an A-curve pattern. ("Neural" 341)

It is not the case that a single feature variant will be the only realization of a linguistic feature for any large group of speakers, whether socially or regionally defined. Instead, features

are either *more* or *less* frequent among various groups, and because of saliency and human perception, those features which are at the top of the A-curve (very frequent) for any sub-sample will be perceived as "normal" and those features which are in the tail of the A-curve (uncommon) for any sub-sample will be perceived as "different" within that sub-sample. This is an important observation which explains how a language can be highly variable and yet still be perceived as a single language: "[T]he top-ranked features may be different at different levels of scale even within the same data set, and so we can account for the universal impression of speakers that there is characteristic language variation between different regional or social groups, at the same time that speakers can have the impression of an overall language at a higher level of scale" (Kretzschmar, "Dialectology" 20).

It is evident even from just a cursory examination of the frequency graphs for these 12 lexical items[15] that there is a definite pattern emerging. While the graphs may vary in minor ways, such as the location of bumps and dips, every single graph is in the form of an A-curve: "an asymptotic curve with a high limit at the Y-axis and a low limit along the X-axis" (Kretzschmar, Linguistics 93). As noted above, though, there are a few varieties of the standard A-curve: Distribution Types A, B, and C.

An A-curve with Distribution Type A is characterized by the presence of a single highly-frequent item, several items with moderate frequencies, and many items with very low frequencies. This distribution results in a curve with three distinct sections: a rapid initial drop, a less abrupt decline, and a long tail. There are sometimes "bumps" present in the middle of these three sections, caused when two or more responses have very similar frequencies in the middle

[15] See Appendix E for a complete collection of the All Responses and subset graphs for each lexical item.

of a rapid decline. The “meadow” All Responses graph illustrates Distribution Type A with no bump in the middle section:

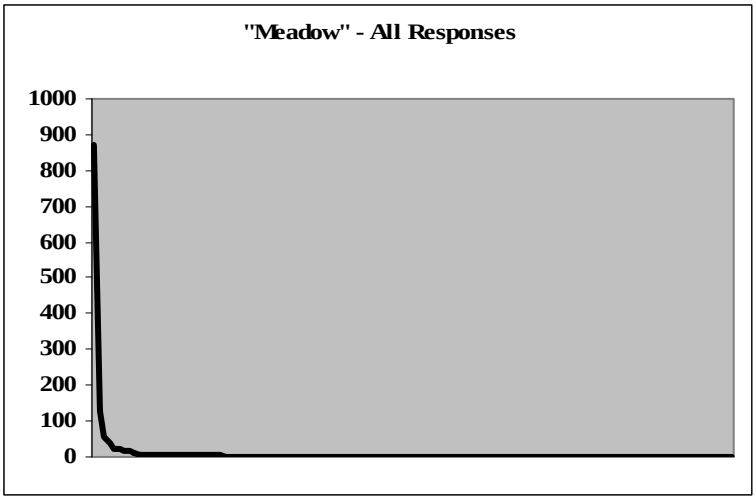


Figure 38: “Meadow” All Responses

The majority of the 108 graphs I examined exhibit Distribution Type A, including all of the “hog pen” graphs except the NY/NJ/PA subset, all of the “meadow” graphs, all of the “swamp” graphs, all of the “corn bread” graphs except the NY/NJ/PA subset, all of the “pancakes” graphs, the “andirons” All Responses graph and half of the subset graphs (Males, Females, Age Group 1, NY/NJ/PA), all of the “hearth” graphs, all of the “pallet” graphs, six of the “cloudburst” graphs (Males, Females, Age Group 3, NY/NJ/PA, MD/DC/DE/VA/WV, NC/SC/GA/FL), three of the “dry spell” subsets (Females, Age Group 1, NC/SC/GA/FL), and all of the “steady drizzle” graphs.

Those graphs which illustrate Distribution Type B are characterized by the presence of three or more responses with very high frequencies, followed by a drop in frequency and the presence of more sets or pairs of items with similar frequencies, a pattern that continues into the tail of the curve. Such a pattern results in a line that gradually curves downward, typically with

many bumps along the way, but with no distinct sections. The only graphs which exhibit this type of distribution are those for the "cobbler" data set; the All Responses graph is shown below.

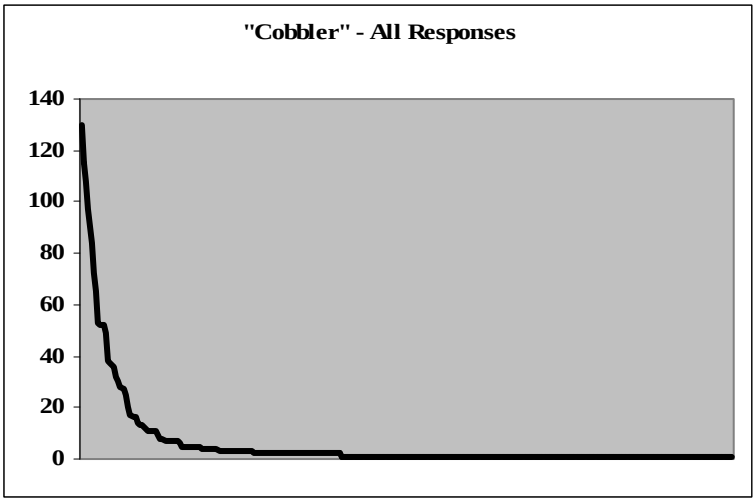


Figure 39: "Cobbler" All Responses

The "cobbler" data set included the largest number of different responses and unique items among all of the data sets; however, the majority of responses provided by the 1,162 respondents centered around three different lexical items, resulting in Distribution Type B. It is possible that other lexical items which were not examined in this study may show the same pattern of distribution, but it is difficult to estimate how prevalent this distribution may be, based on the data in this study alone.

Distribution Type C occurs when the first two items in the data set have approximately equal frequencies, but are followed by items with much lower frequencies. The line for this type of graph begins with a very gradual downward slope, then drops rapidly and ends with the long tail that characterizes all A-curves, as illustrated by the graph below.

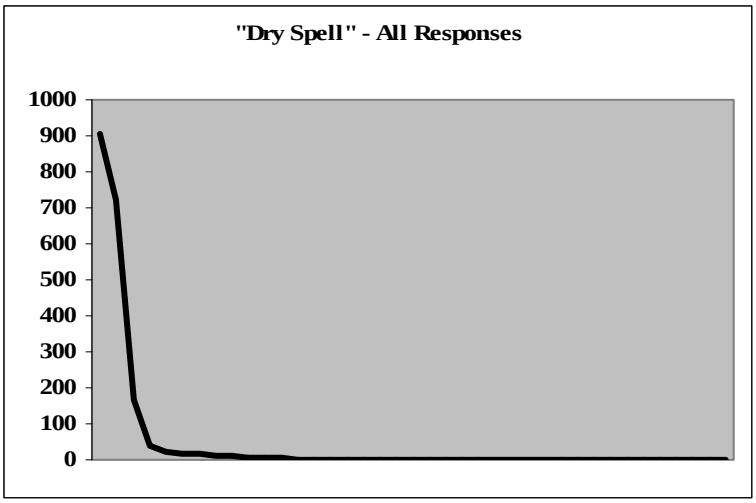


Figure 40: "Dry Spell" All Responses

The graphs which exhibit this type of distribution include the NY/NJ/PA "hog pen" graph, the NY/NJ/PA "corn bread" graph, half of the "andirons" subset graphs (Age Group 2, Age Group 3, MD/DC/DE/VA/WV, NC/SC/GA/FL), three of the "cloudburst" graphs (All Responses, Age Group 1, Age Group 2), and six of the "dry spell" graphs (All Responses, Males, Age Group 2, Age Group 3, NY/NJ/PA, MD/DC/DE/VA/WV). Although most of the Distribution Type C graphs come from the weather domain, all three of the other domains are represented by at least one graph showing this pattern of distribution.

Overall, Distribution Type A is the most common pattern for the A-curves in this examination (84 out of 108 charts—78%). Distribution Type C accounts for 15 of the 108 charts (14%), while Distribution Type B occurs in only 9 of the charts (8%). However, all of the distribution types I have detected are quite similar to one another, differing only in the degree and exact location of frequency decline. The presence of the *basic* A-curve shape in each of the graphs examined above is significant because it proves that while language is highly variable, the variation "always shows the same basic pattern of distribution" (Kretzschmar, Linguistics 93). It is probable that additional analyses of other linguistic features will result in similar results; this

pattern of variation and the presence of the A-curve distribution for linguistic features is assumed to be both stable and systematic (Kretzschmar, “Distributional” 399).

The final feature of language variation to be examined here is the assumption that frequency graphs of linguistic features will exhibit scaling: “aggregation of more and more of the basic patterns, forming new patterns at higher levels of scale” (Kretzschmar, “Dialectology” 9). The majority of the subset graphs for any particular data set are typically scaled-down versions of the All Responses graph for that same data set, although the presence of certain frequency anomalies occasionally results in a graph with a slightly different appearance. The “hog pen” data set is particularly illustrative of this pattern—seven of the eight subsets have frequency graphs that are reminiscent of the All Responses graph, and three of the eight are practically identical to the All Responses graph. In this instance, the only graph that doesn’t seem to exhibit perfect scaling, that of the NY/NJ/PA subset, still follows the same general A-curve shape and the difference can be explained by the relatively small number of unique responses provided. For all of the other “hog pen” graphs, though, the basic shape and proportions of the All Responses curve have remained the same while the actual size of the curve has been reduced in the other subset graphs (see Kretzschmar, “Dialectology” 17 for further explanation and another illustration of this effect). Additionally, while the graphs are very similar to one another, the lexical items in each of the top 10 positions are not always consistent across subsets, as discussed in the previous chapter.

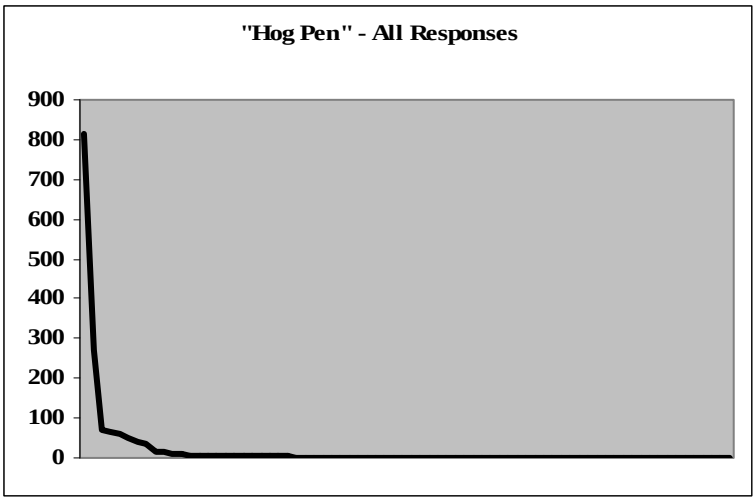


Figure 41: "Hog Pen" All Responses

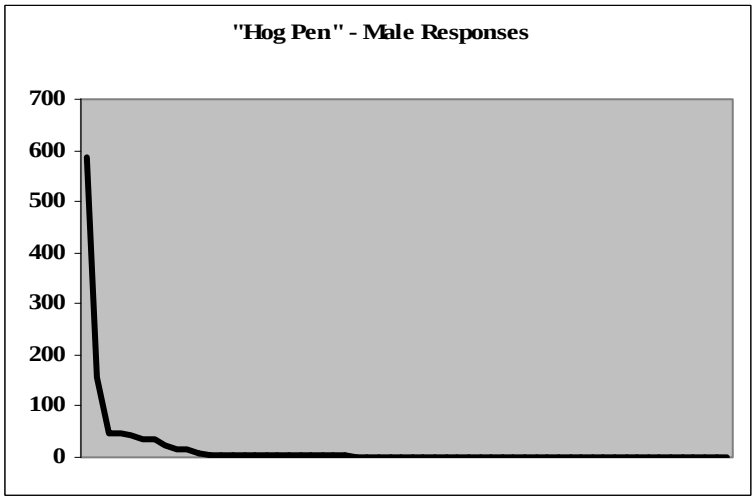


Figure 42: "Hog Pen" Male Responses

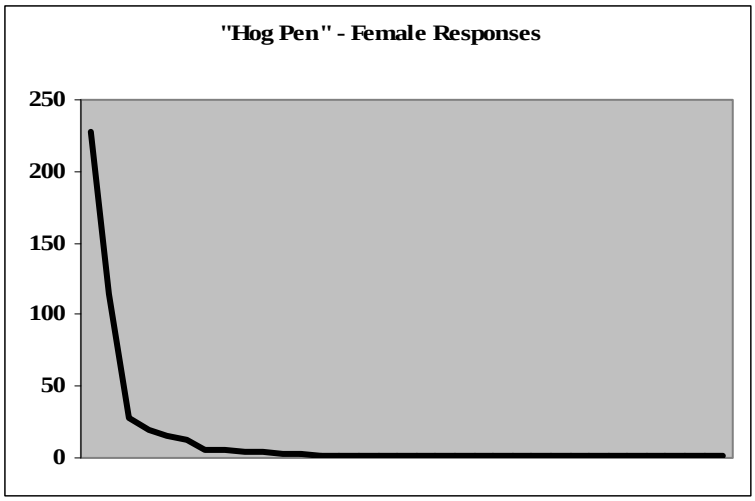


Figure 43: “Hog Pen” Female Responses

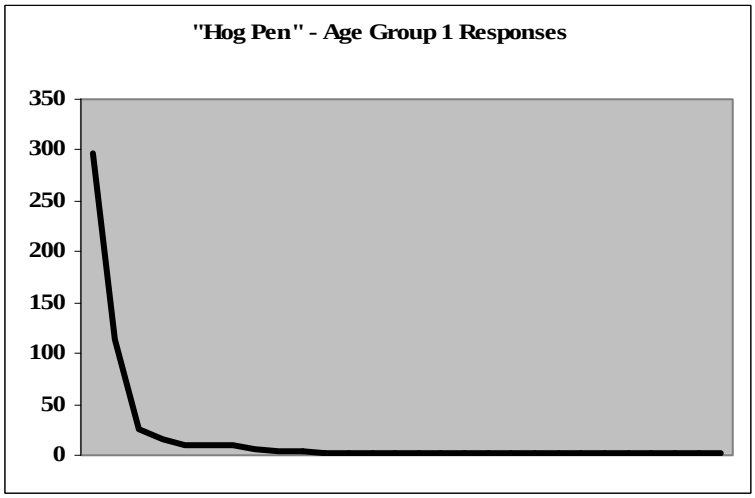


Figure 44: “Hog Pen” Age Group 1 Responses

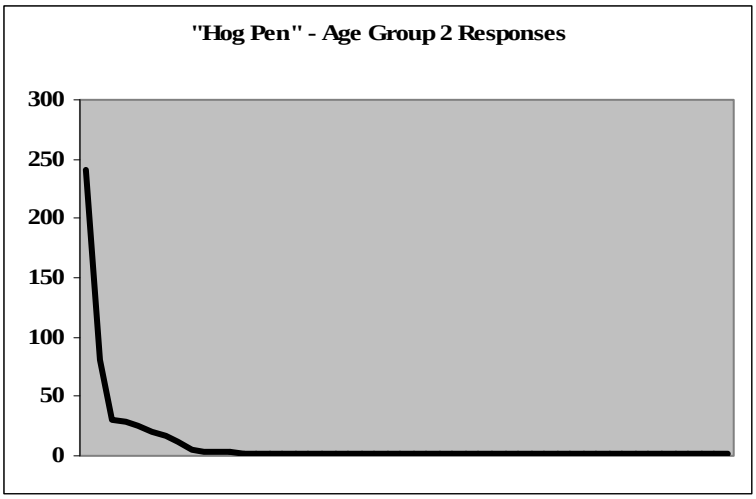


Figure 45: "Hog Pen" Age Group 2 Responses

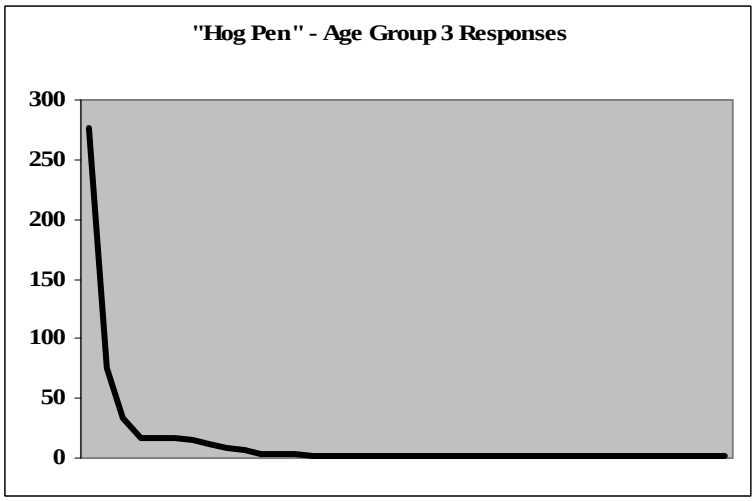


Figure 46: "Hog Pen" Age Group 3 Responses

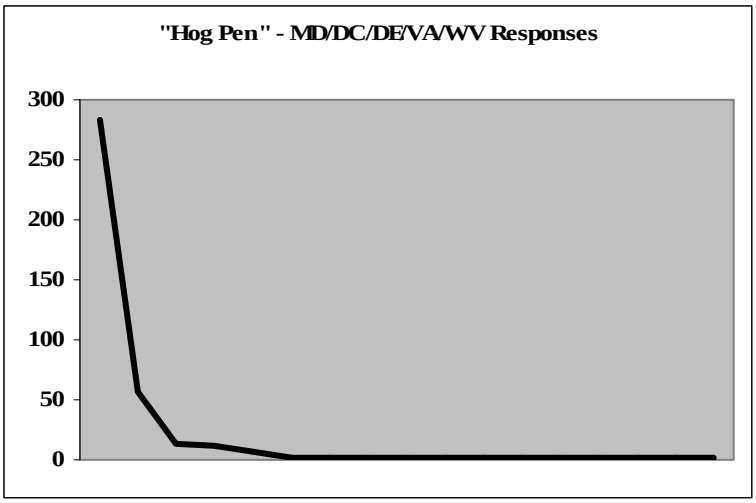


Figure 47: "Hog Pen" MD/DC/DE/VA/WV Responses

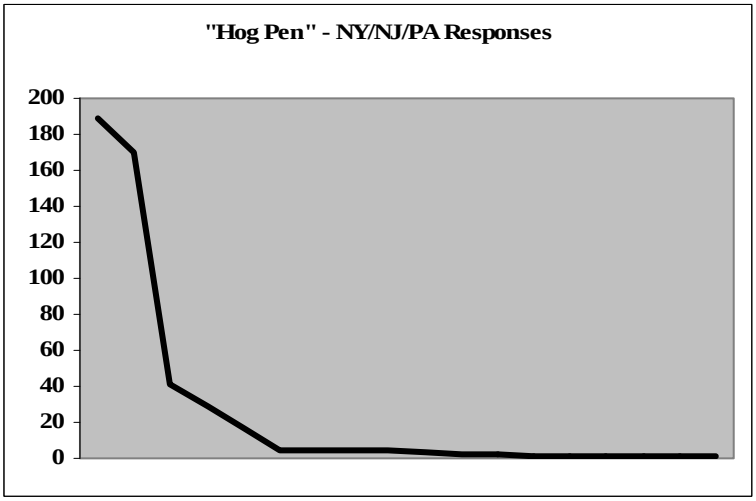


Figure 48: "Hog Pen" NY/NJ/PA Responses

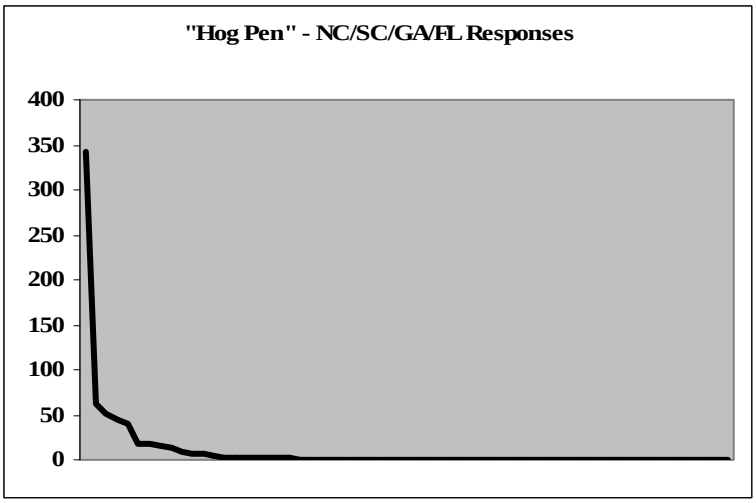


Figure 49: "Hog Pen" NC/SC/GA/FL Responses

The presence of scaling effects is significant because it shows that while language is highly variable, it is not chaotic. Instead, the variation that exists in human language follows certain patterns that may not be readily apparent on the surface. Kretzschmar explains this phenomenon in terms of a metaphor: "We can expect to observe what amounts to an unlimited series of Russian dolls in speech, in which the dolls have the same shape at different scales, but may each be painted with different motifs and colors. The property of scaling tells us…to look for the same patterns composed of different elements at different scales of observation" (Linguistics 184). So, for example, the English language as a whole can be thought of as the largest of the dolls, in which fits the speech of Northerners and Southerners, men and women, young and old speakers. Each of these groups of speakers has a slightly different language, "painted" with different colors and patterns, but the basic "shape" of each is identical.

CHAPTER 6
CONCLUSION

Language variation is highly systematic and regular, as demonstrated by this examination of LAMSAS lexical item survey data. Furthermore, it exhibits characteristics of complex systems, such as the emergence of order and the presence of scaling. Both of these findings are of profound consequence for the study of language variation and change, fields which, until recently, have been dominated by structuralist perspectives. While Saussure, Chomsky, and their respective students have provided useful theories with which to analyze language, the linguistics of speech offers a valuable new model to those linguists who wish to study language variation and change from a fresh perspective. For example, Kretzschmar explains that the A-curve provides an appealing new way of analyzing language change over time:

> Since we hypothesize that an A-curve will exist for every feature at any moment in time (i.e., that language will not suddenly become invariant), we can define the notion "linguistic change" itself as the change in the location of the target variant at different heights along the curve. If a particular variant occurs at a higher place on the curve than it did before, it has become more frequent and so we can say that the direction of change for that variant is positive; if a variant occurs at a lower place on the curve than it did before, it has become less frequent and the direction of change is negative. (Kretzschmar, "Distributional" 394)

This approach to the study of language change over time is illustrated with the following set of graphs, from Kretzschmar's "Distributional Foundations for Language Change". These

graphs trace the location of a specific variant on a frequency chart over time; in this instance, the variant has become more frequent and thus illustrates a positive direction of change.

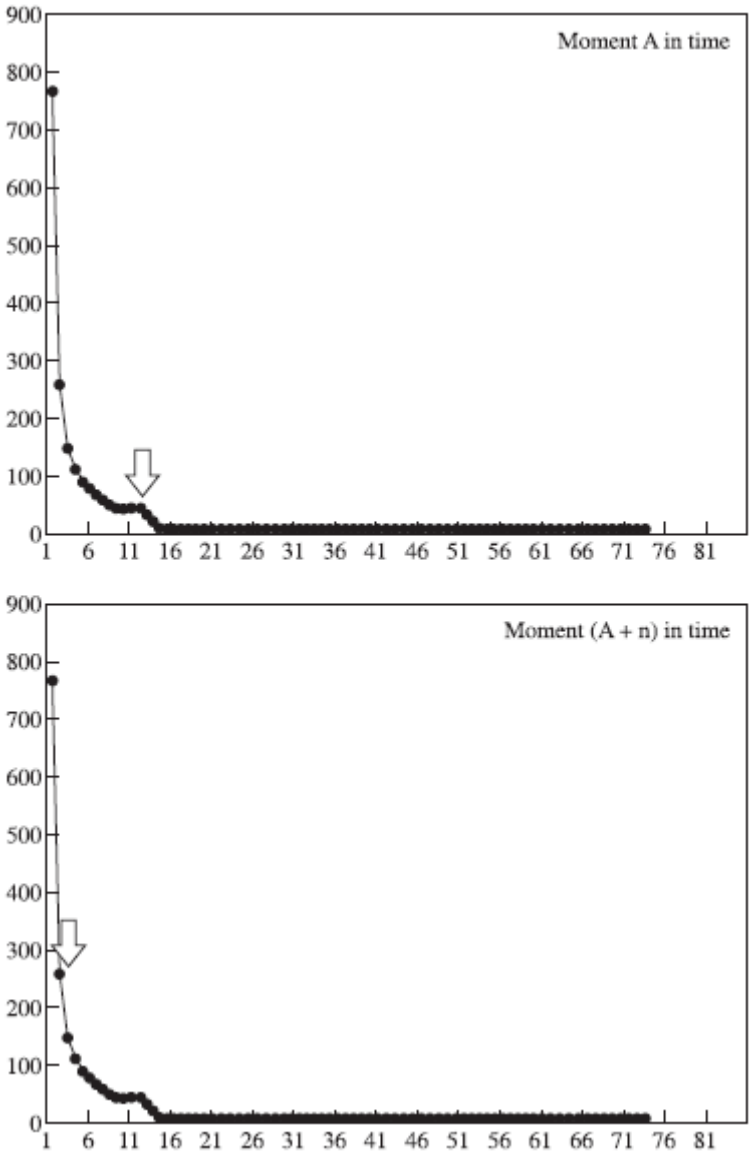


Figure 50: An A-curve at two different moments in time (Kretzschmar, "Distributional" 395).

The study of language change over time may become even more important in the future, despite the popular belief that language is becoming less variable over time. Ellen Johnson conducted a 1990 survey similar to that of the LAMSAS surveys from the 1930s, although on a

much smaller scale. Her surveys not only show "evidence of lexical change," but also the perpetuation of "core terms" (Burkette 140): those terms which were highly frequent in the 1930s data. Additionally, Allison Burkette's 1998 examination of the lexical item "bureau" produced "the same kind of variation found in linguistic atlas data. For each picture there were core and peripheral terms used to identify the visual image—the core terms eliciting the highest number of responses, the peripheral terms given much less frequently" (141). Finally, in her 2001 examination of change in farming terms from Linguistic Atlas data, Sandra Hoover explains, "[T]he twenty target lexical items produced almost eight variants apiece for the LAGS [Linguistic Atlas of the Gulf States] data and almost thirteen variants apiece for the 2001 data. This amount of variation confirms Johnson's findings that there is rich lexical variety in the South and that the lexical variety is increasing" (51). Although the LAMSAS survey data I used for this study is a valuable resource, it is vital that linguists continue to conduct similar surveys and collect large amounts of linguistic data. Continuous examination of language variation and change will no doubt result in additional insights and further adaptations of the model of the linguistics of speech.

The remarkable amount of variation within human languages means that "all languages and dialects are 'endangered' at all times by constant linguistic change and cultural change, in the sense that no language or dialect can stay exactly the same as time passes" (Kretzschmar, Linguistics 20). Language is highly variable because people are always making choices (often unconsciously) about what they say or how they say it. Although human languages are constantly shifting and are not static, speech is "not chaotic or unmanageable, but rather offers regularities across the continuum that the linguist can use to address particular problems having to do with language" (Kretzschmar, Linguistics 52). Human language, in its many various

forms, conforms to the principles of complex systems and therefore can be analyzed according to the linguistics of speech model.

APPENDIX A
DELETED RESPONSES FOR EACH LEXICAL ITEM

Cobbler—263 deleted responses:

NR (247 occurrences); a ?? pudding; and good they are too; ? pies; blackberry fl??; crap???; crust is bottom and top; just two thick pies together; lattice covered over first; r???per???; rollerr pollers; rancid; was thick; washes around; y[16]; down in bottom; raisin leaves

Pancakes—85 deleted responses:

darkies called them; darkies use to call them; hamm?? gl ds???; iron to make them; kicklen day; my favorite; nigger calls them; NR (72 occurrences); old term; olie koek; really thin; pannekoek; sop them with sugar; suppose to be sweet

Cornbread—77 deleted responses:

and water; at the fire; buelen; carm??n corn bread; find a trial; fried in a pan of grease; he lit the pipe; meal milk eggs soda salt; nigger hill; NR (57 occurrences); pone ?; pone of; Sally Lunn; salt; saleratus; soda; egg (2 occurrences); eggs; milk (2 occurrences)

Meadow—191 deleted responses:

-0- (6 occurrences); cow'll bog in them; cut it for hay; eine Wiese (2 occurrences); for pasture; has water settled; NA (13 occurrences); no trees don't grow; NR (161 occurrences); off of it; or; plaatsje; Wiese

Hog pen—69 deleted responses:

??? (2 occurrences); -0-; NA (11 occurrences); NR (55 occurrences)

Swamp—84 deleted responses:

-0- (15 occurrences); Begin Swamp; Cannawanga Swamp; can't; Carver's Bay (2 occurrences); Copper Swamp; drain the swamp; Elm Lake; Florida; got no bottom; Hal's Pond; Hell Hole Bay; Hellhole Swamp; hogs ranged; Johnson's Swamp; Little Salkehatchie Swamp; Montezuma Marshes; Murray's Swamp; NA (13 occurrences); not; not tenable; NR (19 occurrences); Oak Swamp; Okefenokee (2 occurrences); Okefenokee Swamp (2 occurrences); Pale Pond; Piney Bay; Pullman's Lake; Princess Pond; Raper Bay; Rock Hole Swamp; Salt Catcher Swamp; seemingly has no bottom; Swamp Rabbit; The Bog on Vly; titi; Vly Mountain; Barren Swamp

Cloudburst—116 deleted responses:

-0-; go over kind of quick; godsend; heavy weather that make up and; it's steady; in floods; it poured in torrents; NA (62 occurrences); NR (45 occurrences); wouldn't that soak you; sky

[16] There were two responses listed for the same respondent: "brown bettu" and "y". These two were combined into the responses "brown betty," which was a common response for other respondents.

Dry Spell—89 deleted responses:

ein trockne Spell; going to get dry enough to; it’s some kind of dry; longer; month; NA (6 occurrences); nearly every vegetation dies; NR (71 occurrences); often; parching the corn; some says; two weeks of clear weather; burn up the crops

Steady drizzle—133 deleted responses:

-0- (3 occurrences); best rain; brief; a steady drizzle ain't nothing much; enough to settle the dust; even the steady steady drizzle; holds on lasts longer; if the sun's a-shining it'll rain tomorrow; it lasts a long time; it's steady; just did drop; lasts; light rains is a heap better on the crops; NA (63 occurrences); NR (53 occurrences); to lay the dust; used to call

Hearth—28 deleted responses

colored people called it; Feuerherd; fire in; fire in it; four; hearth to it; in the fireplace; is right; isn't it; NA (5 occurrences); NR (12 occurrences); niggers call it; you doesn't sweep it out after sunset

Pallet—322 deleted responses

cama redonda; NA (14 occurrences); NR (305 occurrences); Sprausack; under there

Andirons—114 deleted responses

-0- (3 occurrences); Feuerhund; made by an old-timey blacksmith; NA (6 occurrences); NR (99 responses); out of iron; sitting up on prong; way back in time in my growing up; you lay logs on

APPENDIX B
COMBINED RESPONSES FOR EACH LEXICAL ITEM

Cobbler:

brown bettu/y = brown betty for same respondent; berry pies + a berry pie; a betty + betty; a crows nest + crows nest; a deep apple pie + deep apple pie; a deep pie + deep pie + deep pies; a pie + pie + pies + them pies; a pot pie + pot pie + pot pies; an apple turnover + apple turnover; apple cobbler + apple cobblers; apple dumpling + apple dumplings+ apple dumping; apple jack/s + applejack; apple pie + apple pies + the apple pie; apple pot pie + apple pot pies; apple pudding + apple puddings; apple tart/s; apple john + apple-john/s; apple/s; baked apple/s; big pie/s; birds nest + bird's nest; bird/s nest pudding; birrds nest + birds nest; blackberry/ies; blackberry dumpling/s; blackberry pie/s; cherry pie/s; cobbler/s; cobbler/s pie; custard/s; dumpling/s; dutch apple pie/s; floating isl+/s; fruit pie/s; huckleberry pie/s; lemon pie/s; minced meat pie/s; minced pie/s; pan pie/s; peach cobbler/s; peach pie/s; pear pie/s ; pecan pie/s; plate pie/s; potato custard/s; potato pie/s; pudding/s; pumpkin pie/s; sanker pie/s; sanker/s; tart/s; tater/tator pies; turnover/s; washington pie/s

Pancakes:

batter bread/s; batter cake/s + battercake/s + battercakes; buckwheat cake/s; cookie/s; corn fritter/s; corn cake/corncakes; flannel cake/s + flannel cale; flannen cakes + flannen cakes; flapjack/s; flitter cake/s; flitter/s + flitters (bread); flour cakes + flourcakes; fried cake/s; fry bread/s; griddle cake/s; griddle/s; hoecake/s; hotcake/s; pancake/s; puff/s; slapjack/s; tea cakes/teacakes; wheat cake/s + wheatcakes

Cornbread:

a/pone of bread; a pone of corn bread + pone of corn bread + pone of cornbread; ash cake/s; corn bread loafs/loaves; corn cake/s; corn dodger/s; corn loaf/loaves; corn meal muffin/s; corn muffin/s; corn pone/s + corn pone (2 spaces between “corn” and “pone”); corn bread + cornbread; corn meal bread + cornmeal bread; corn meal muffins + cornmeal muffins; hoe cake corn bread + hoecake corn bread; hoe cake/s + hoecake/s; hush puppies + hushpuppies; johnny cake/s + johnnycake; loaf + loaves; muffin/s; plain corn bread + plan corn bread; steam/ed corn bread; steam/ed bread; pone/s of corn bread; pone/s of bread

Meadow:

bog hole/s; bog/s; bottom/s; canebreak/s; flat/s; grass bottom/s; hayfield/s; marsh/es; meadow/s + The Meadow; meadow bottom/s; meadow land/s; pasture/s; pasture land/s; pine barren/s; pond/s; savanna land/s; savanna/s+ the old savanna; swale/s + down in the swale

Hog pen:

fattening pen/s; floored pen/s; hog crawl/s; hog lot/s; hog pasture/s; hogpen + hog pen/s; hog sty/ies; pig pen/s

Swamp:

all boggy + boggy; basin/s; bayou/s; bay/s; bog/s; bog hole/s; bog land + bogland; boggy place/s; Boggy Pond + boggy pond; cow mire/s; fresh water marsh/es; hammock/s; highland + high land; interval/e; island/s; knoll/s; lake/s; low place/s; lowland/s; marsh land + marshland; marsh/es; marshy place/s; mill pond/s; mud hole/s; overflow + overflowed + overflown; pond/s; river swamp/s; river/s; savanna/s; slash/es; Slough + slough/s; swamp land/s; swamp/s; swampy land/s; swampy place/s; tussock/s

Cloudburst:

good rain + a good rain; a pour + pour; a very heavy rain + very heavy rain; a/washing rain; big rain/s; cloudburst/s; downpour/s ; freshet/s; hard rain/s; lighterd knot floater/s; raining bullfrog/s; shower/s; soaker/s; stump lifter/s; torrent/s; trash lifter/s; trash mover/s; trash piler/s; pour out + pourout; rained/raining cats and dogs

Dry Spell:

one little dry spell + dry spell/s; drouth/s; a/long dry spell

Steady drizzle:

devil's beating his wife/behind the door; steady drizzle + steady drizzled + steady drizzles + drizzling + it steady drizzled + just a steady drizzle; fodder shower/s; steady drizzle all day + it steady drizzle all day long; steady drizzle/d all day + it steady drizzled all day; drizzly + just drizzly; light shower/s; little shower/s; mist + misting; rainy day + just a rainy day; set rain/s; settle the dust + settling the dust; shower + showered + showering + showers; sprinkle + just a sprinkle + sprinkled + sprinkling + sprinkles; steady + just steady; summer shower/s

Hearth:

fireplace + fireplaces; fire + fires; grate + grates; hearth + hearths; hearth rock/s + hearthrock/s

Pallet:

bed/s; featherbed/s; mattress/es; pad/s; pallet/s; straw tick/s; trundle bed/s

Andirons:

andiron/s; fire horse + firehorse; handiron/s; dog/s; dog iron/s; fire log/s; firedog/s; grate/s; iron/s; fender/s; fire rock/s; railroad iron/s; rock/s; wood dog/s

APPENDIX C
AGE GROUP DATA FOR EACH LEXICAL ITEM

Cobbler: Deleted 2 responses because of "0" age (birds nest, yankee pie)
569 responses per group
- Age Group 1: Ages 16-62
- Age Group 2: Ages 62-75
- Age Group 3: Ages 75-95

Pancakes: Deleted 1 response because of "0" age (pancakes)
886 responses per group
- Age Group 1: Ages 16-59
- Age Group 2: Ages 59-74
- Age Group 3: Ages 74-95

Cornbread: Deleted 2 responses because of "0" age (corn bread and johnny cake)
617 responses per group
- Age Group 1: Ages 16-60
- Age Group 2: Ages 60-75
- Age Group 3: Ages 75-100

Meadow: Deleted 4 responses because of "0" age (meadow, meadow, meadow, swale)
460 responses per group
- Age Group 1: Ages 16-61
- Age Group 2: Ages 61-75
- Age Group 3: Ages 75-95

Hog pen: Deleted 5 responses because of "0" age (fattening pens, floored pens, hog pen, pastures, pig pen)
516 responses per group
- Age Group 1: Ages 16-60
- Age Group 2: Ages 60-75
- Age Group 3: Ages 75-95

Swamp: Deleted 3 responses because of "0" age (river swamp, swamp, swamp)
667 responses per group
- Age Group 1: Ages 16-63
- Age Group 2: Ages 63-75
- Age Group 3: Ages 75-95

Cloudburst: Deleted 3 responses because of "0" age (cloudburst, gully washer and raining pitchfork)

645 responses per group

- Age Group 1: Ages 16-61
- Age Group 2: Ages 61-75
- Age Group 3: Ages 75-95

Dry Spell: Deleted 6 responses because of "0" age (drought, drought, drought, drouth, drouth, drouth)

649 responses per group

- Age Group 1: Ages 16-58
- Age Group 2: Ages 58-74
- Age Group 3: Ages 74-95

Steady drizzle: Deleted 3 responses because of "0" Age (steady drizzle, shower, sprinkle)

622 responses per group

- Age Group 1: Ages 16-62
- Age Group 2: Ages 62-75
- Age Group 3: Ages 75-100

Hearth: Deleted 6 responses because of "0" age (hearth, hearth, hearth, hearth, hearth fireplaces)

571 responses per group

- Age Group 1: Ages 16-62
- Age Group 2: Ages 62-75
- Age Group 3: Ages 75-100

Pallet: Deleted 1 response because of "0" age (pallet)

336 responses per group

- Age Group 1: Ages 16-60
- Age Group 2: Ages 60-75
- Age Group 3: Ages 75-95

Andirons: Deleted 6 responses because of "0" age (andiron, andirons, andirons, andirons, dog irons, firedogs)

510 responses per group

- Age Group 1: Ages 16-59
- Age Group 2: Ages 59-75
- Age Group 3: Ages 75-100

APPENDIX D
STATE DATA FOR EACH LEXICAL ITEM

Cobbler: Average of 570 per group

- New York (307), New Jersey (27), Pennsylvania (147)—**481 total**
- West Virginia (127), Maryland (48), DC (5), Virginia (129), Delaware (9)—**318 total**
- North Carolina (167), South Carolina (413), Georgia (298), Florida (33)—**911 total**

Pancakes: Average of 887 per group

- New York (367), New Jersey (95), Pennsylvania (302)—**764 total**
- West Virginia (208), Maryland (182), DC (8), Virginia (346), Delaware (38)—**782 total**
- North Carolina (317), South Carolina (523), Georgia (249), Florida (25)—**1114 total**

Cornbread: Average of 618 per group

- New York (276), New Jersey (64), Pennsylvania (199)—**539 total**
- West Virginia (157), Maryland (99), DC (2), Virginia (238), Delaware (13)—**509 total**
- North Carolina (204), South Carolina (328), Georgia (252), Florida (21)—**805 total**

Meadow: Average of 462 per group

- New York (355), New Jersey (50), Pennsylvania (182)—**587 total**
- West Virginia (111), Maryland (67), DC (2), Virginia (142), Delaware (18)—**340 total**
- North Carolina (121), South Carolina (214), Georgia (107), Florida (16)—**458 total**

Hog pen: Average of 518 per group

- New York (236), New Jersey (54), Pennsylvania (185)—**475 total**
- West Virginia (117), Maryland (71), DC (2), Virginia (180), Delaware (17)—**387 total**
- North Carolina (174), South Carolina (321), Georgia (183), Florida (14)—**692 total**

Swamp: Average of 668 per group

- New York (312), New Jersey (58), Pennsylvania (168)—**538 total**
- Maryland (71), DC (3), Delaware (16), Virginia (233), West Virginia (115)—**438 total**
- North Carolina (220), South Carolina (488), Georgia (287), Florida (34)—**1029 total**

Cloudburst: Average of 646 per group

- New York (257), New Jersey (59), Pennsylvania (170)—**486 total**
- West Virginia (116), Maryland (90), DC (2), Virginia (214), Delaware (19)—**441 total**
- North Carolina (186), South Carolina (509), Georgia (293), Florida (23)—**1011 total**

Dry Spell: Average of 651 per group

- New York (295), New Jersey (79), Pennsylvania (277)—**651 total**

- West Virginia (211), Maryland (93), DC (2), Virginia (254), Delaware (20)—**580 total**
- North Carolina (216), South Carolina (287), Georgia (202), Florida (17)—**722 total**

Steady drizzle: Average of 623 per group

- New York (363), New Jersey (47), Pennsylvania (152)—**562 total**
- West Virginia (108), Maryland (66), DC (4), Virginia (155), Delaware (16)—**349 total**
- North Carolina (172), South Carolina (474), Georgia (287), Florida (26)—**959 total**

Hearth: Average of 573 per group

- New York (270), New Jersey (51), Pennsylvania (158)—**479 total**
- West Virginia (118), Maryland (69), DC (2), Virginia (199), Delaware (15)—**403 total**
- North Carolina (231), South Carolina (355), Georgia (228), Florida (23)—**837 total**

Pallet: Average of 336 per group

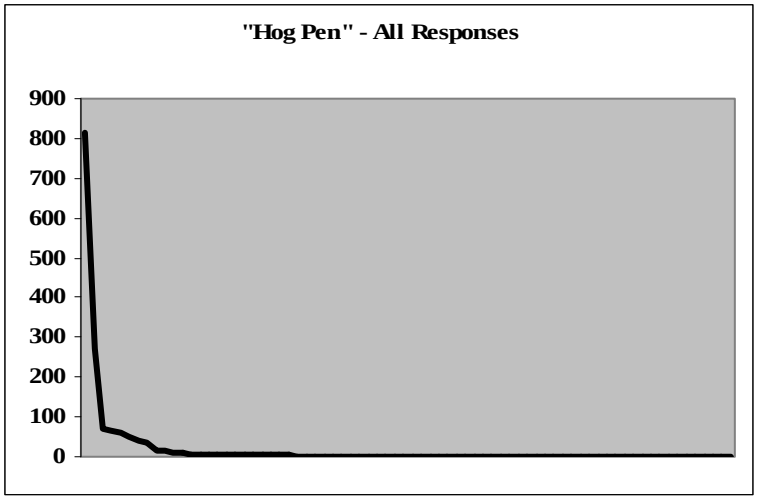
- New York (43), New Jersey (15), Pennsylvania (80)—**138 total**
- West Virginia (103), Maryland (52), DC (2), Virginia (157), Delaware (6)—**320 total**
- North Carolina (188), South Carolina (203), Georgia (145), Florida (15)—**551 total**

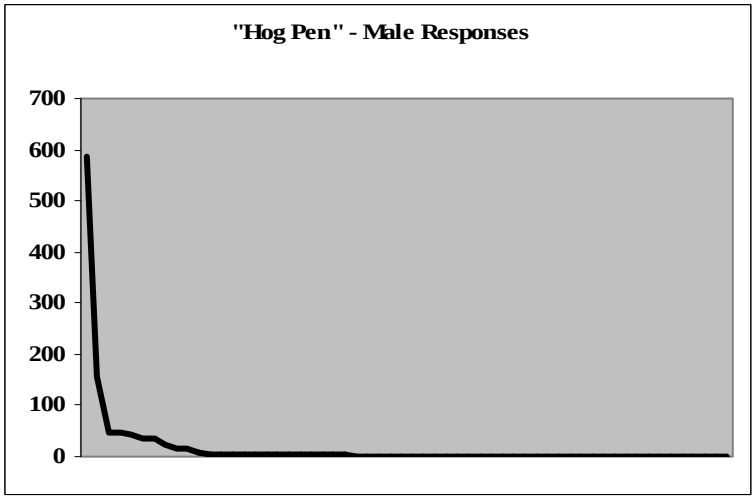
Andirons: Average of 512 per group

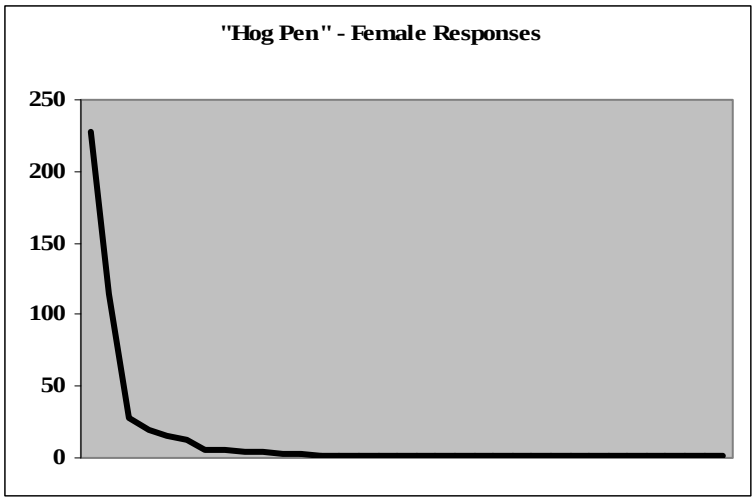
- New York (197), New Jersey (43), Pennsylvania (136)—**376 total**
- West Virginia (132), Maryland (68), DC (2), Virginia (217), Delaware (18)—**437 total**
- North Carolina (218), South Carolina (311), Georgia (179), Florida (16)—**724 total**

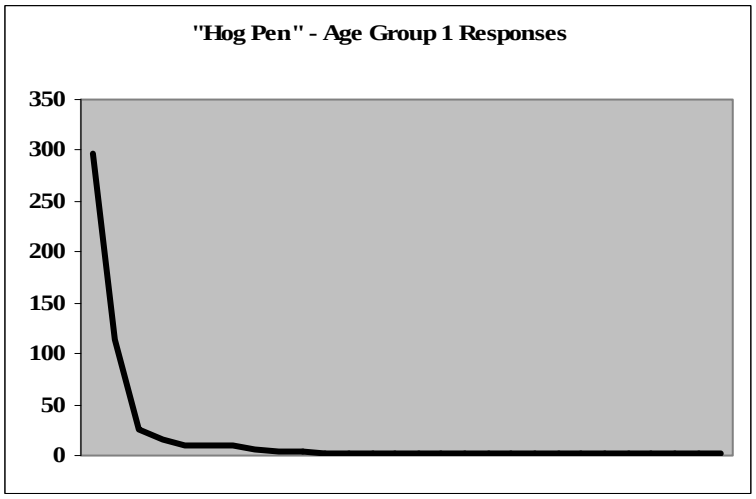
APPENDIX E
FREQUENCY GRAPHS FOR EACH LEXICAL ITEM

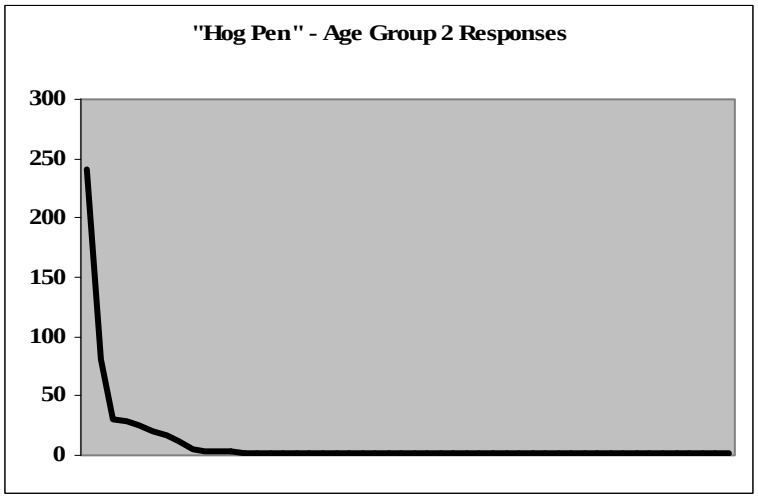
“Hog Pen” Frequency Graphs

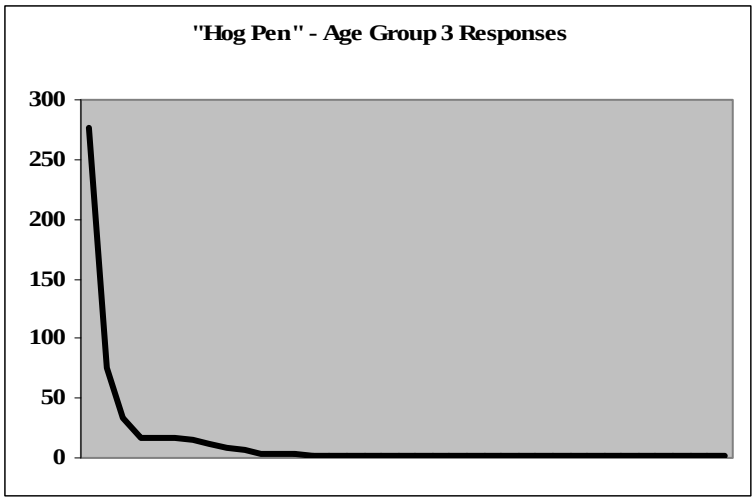


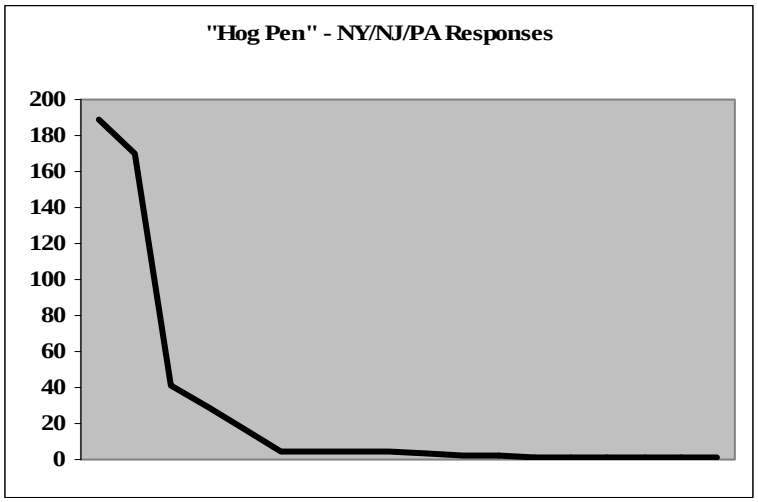


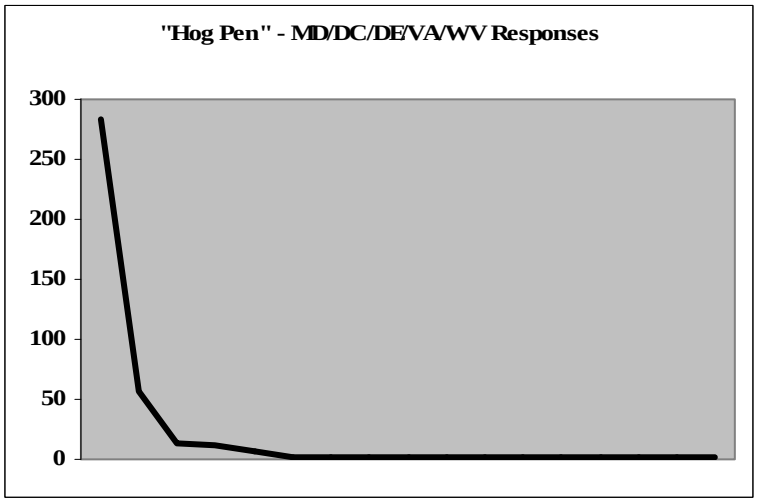

"Hog Pen" - Female Responses
250
200
150
100
50
0

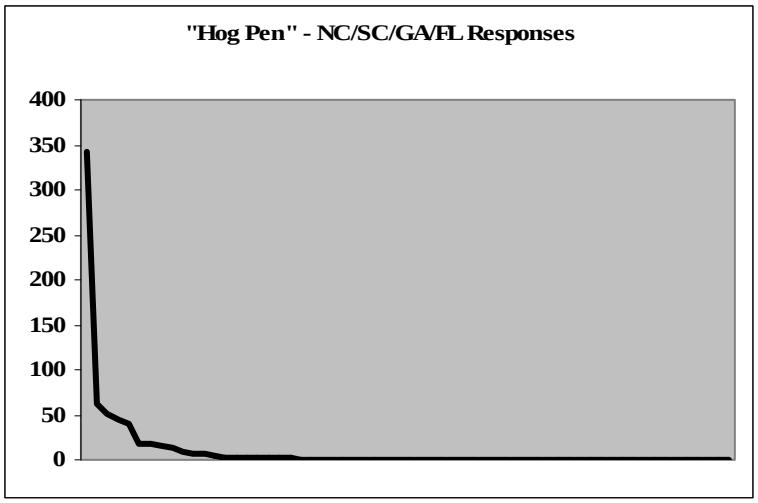

"Hog Pen" - Age Group 1 Responses
350
300
250
200
150
100
50
0


"Hog Pen" - Age Group 2 Responses
300
250
200
150
100
50
0

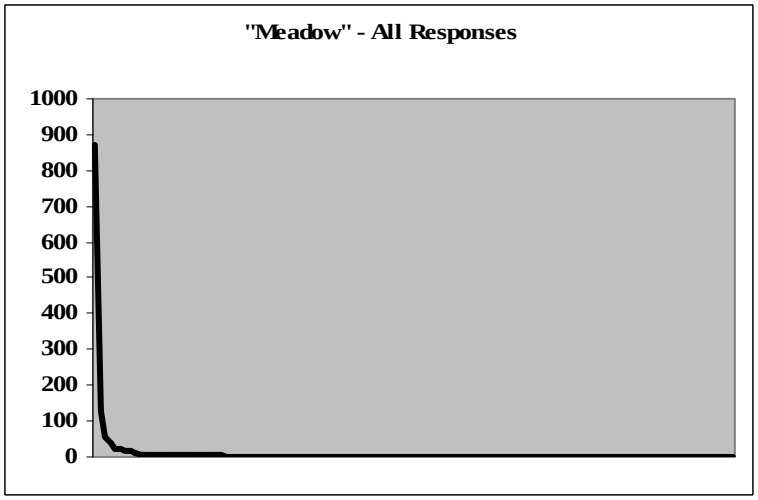

"Hog Pen" - Age Group 3 Responses
300
250
200
150
100
50
0

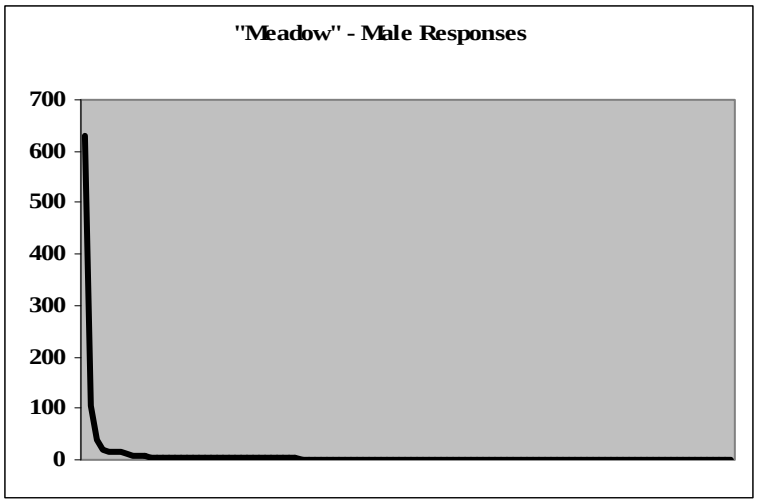

"Hog Pen" - NY/NJ/PA Responses
200
180
160
140
120
100
80
60
40
20
0

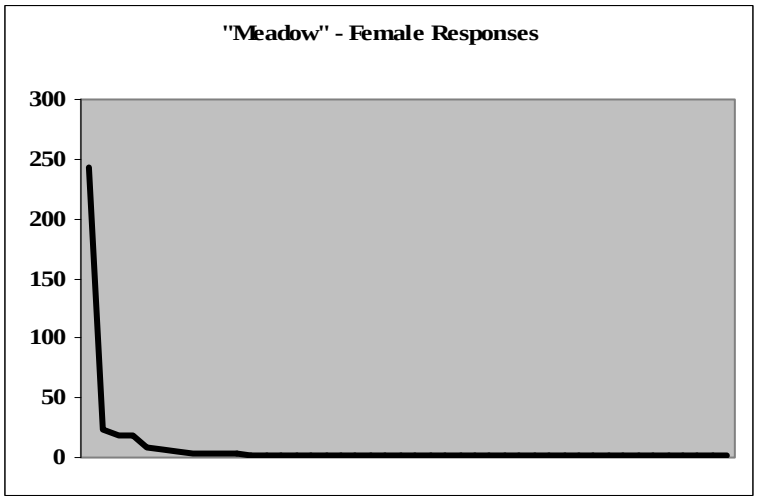

"Hog Pen" - MD/DC/DE/VA/WV Responses
300
250
200
150
100
50
0

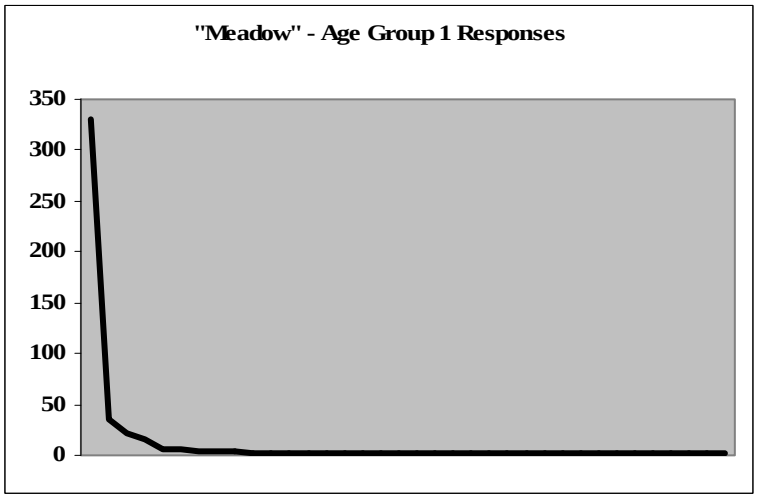


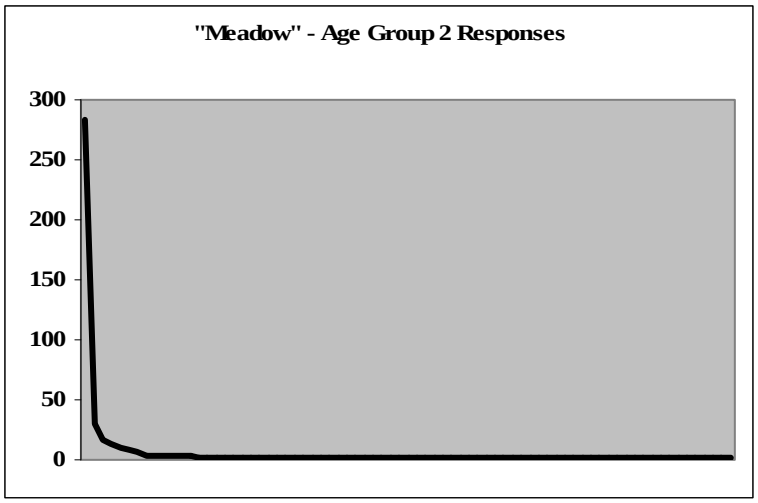
“Meadow” Frequency Graphs

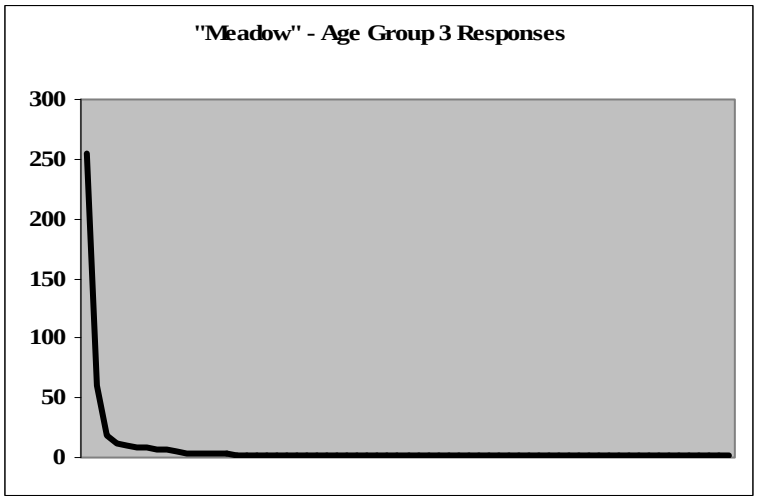


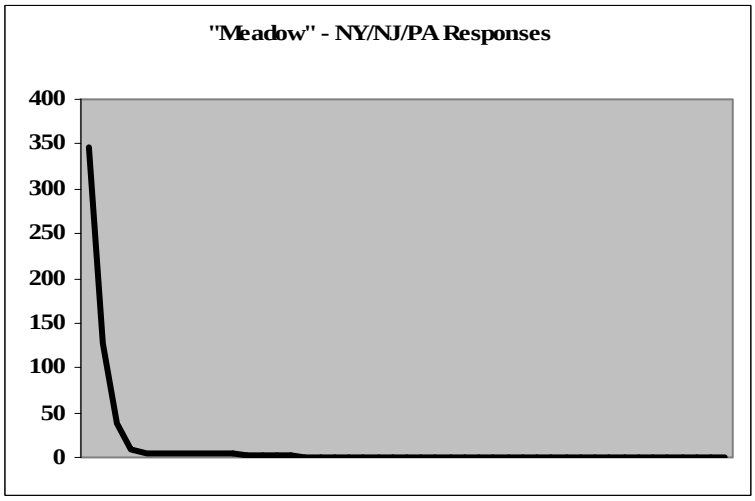

"Meadow" - Male Responses
700
600
500
400
300
200
100
0

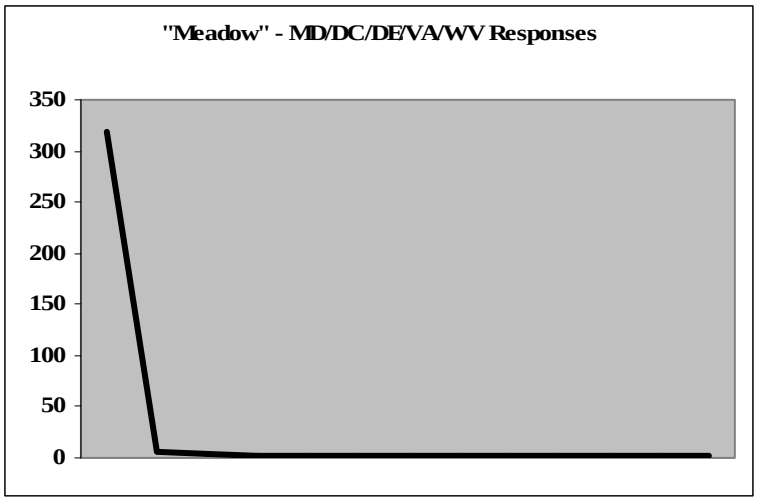

"Meadow" - Female Responses
300
250
200
150
100
50
0

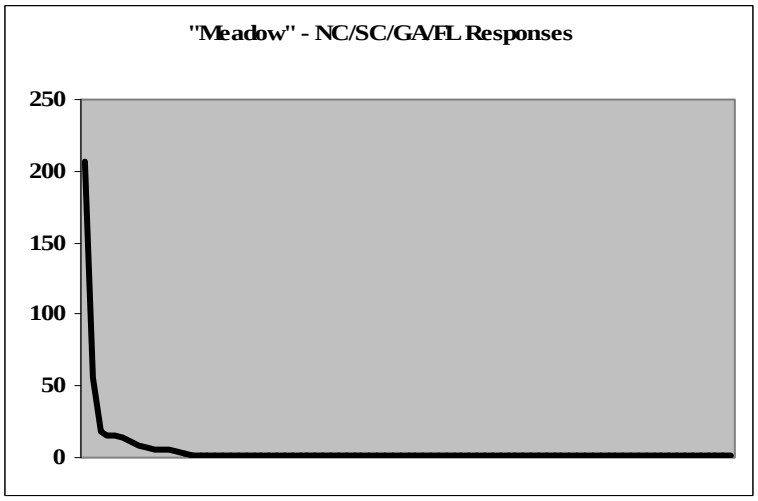

"Meadow" - Age Group 1 Responses
350
300
250
200
150
100
50
0


"Meadow" - Age Group 2 Responses
300
250
200
150
100
50
0

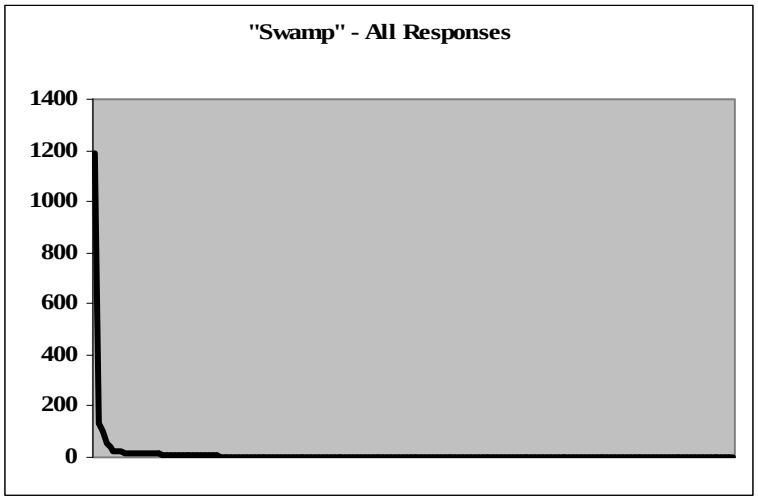

"Meadow" - Age Group 3 Responses
300
250
200
150
100
50
0

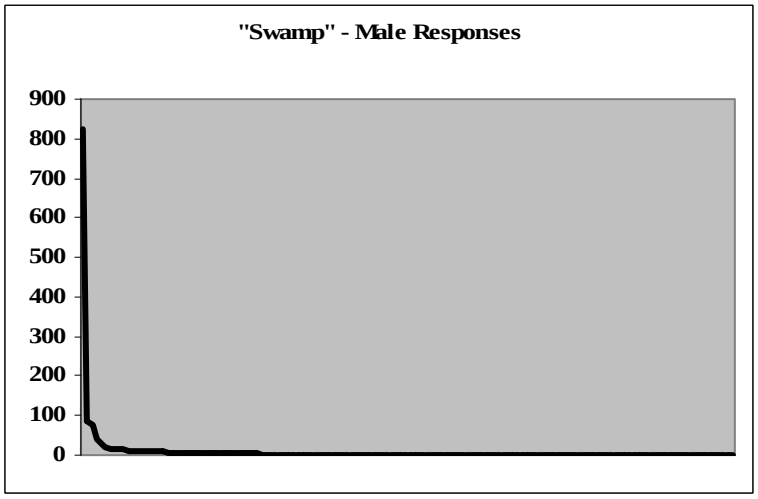

"Meadow" - NY/NJ/PA Responses
400
350
300
250
200
150
100
50
0

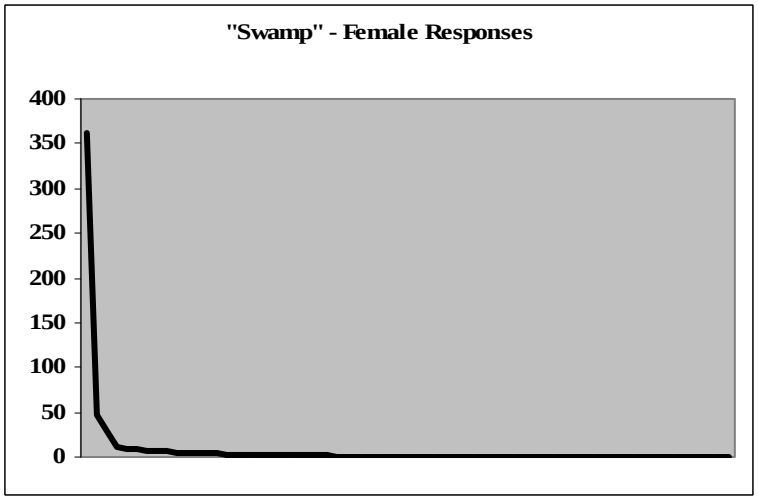

"Meadow" - MD/DC/DE/VA/WV Responses
350
300
250
200
150
100
50
0

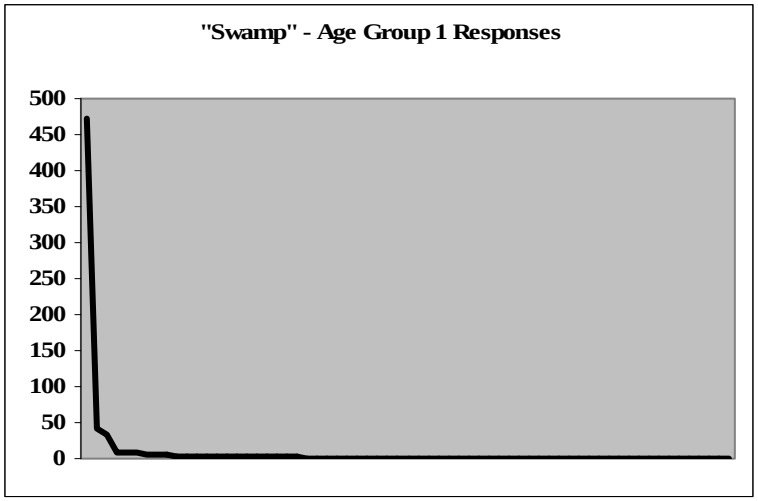

"Meadow" - NC/SC/GA/FL Responses
250
200
150
100
50
0

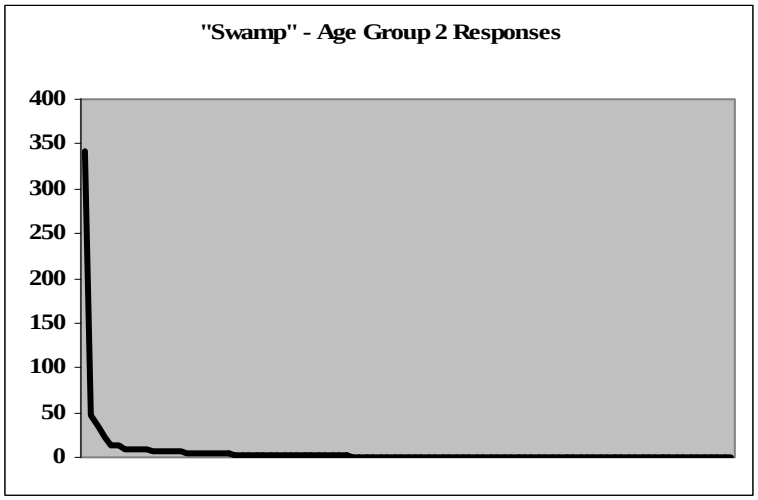
“Swamp” Frequency Graphs

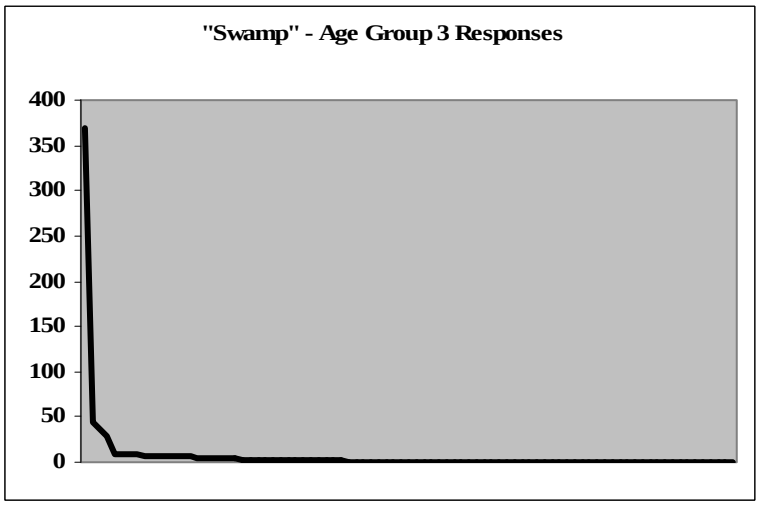


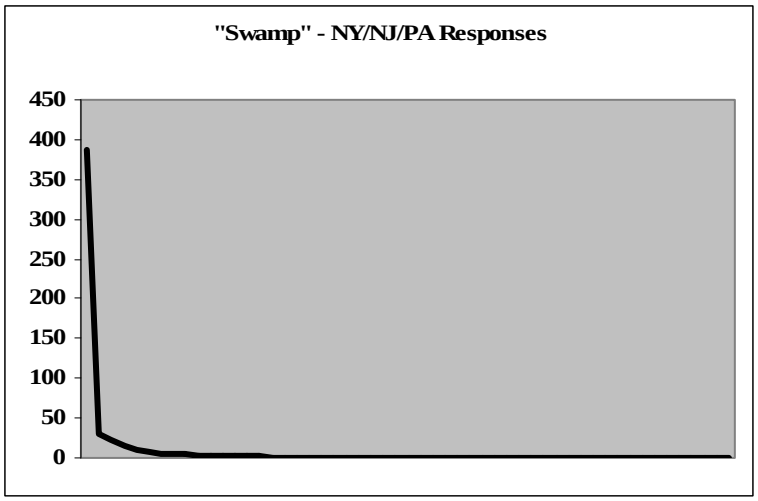


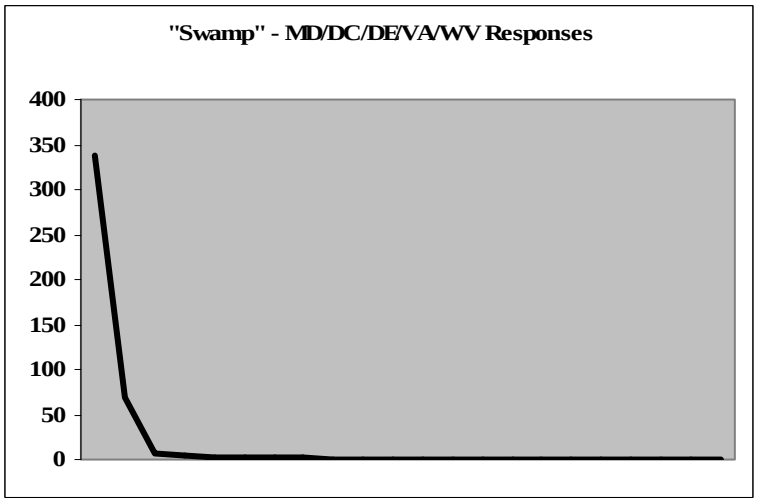


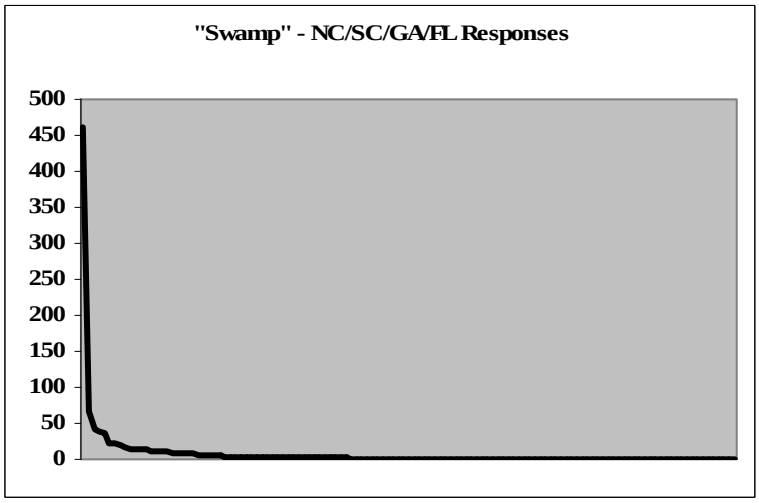

"Swamp" - Age Group 1 Responses
500
450
400
350
300
250
200
150
100
50
0


"Swamp" - Age Group 2 Responses
400
350
300
250
200
150
100
50
0

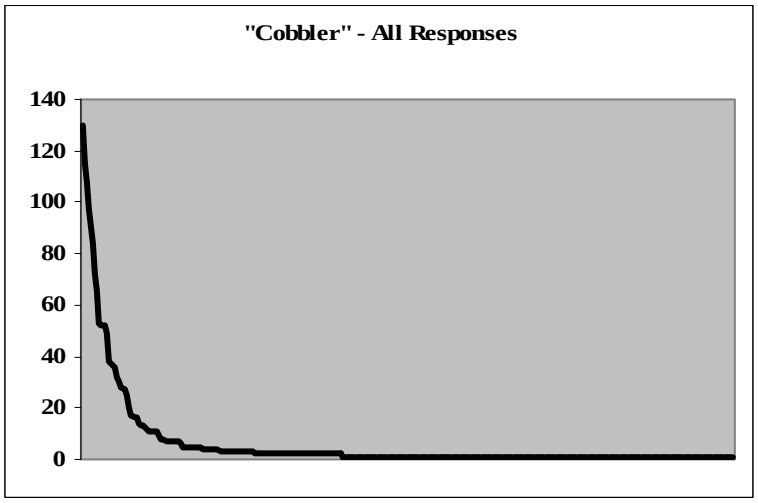

"Swamp" - Age Group 3 Responses
400
350
300
250
200
150
100
50
0

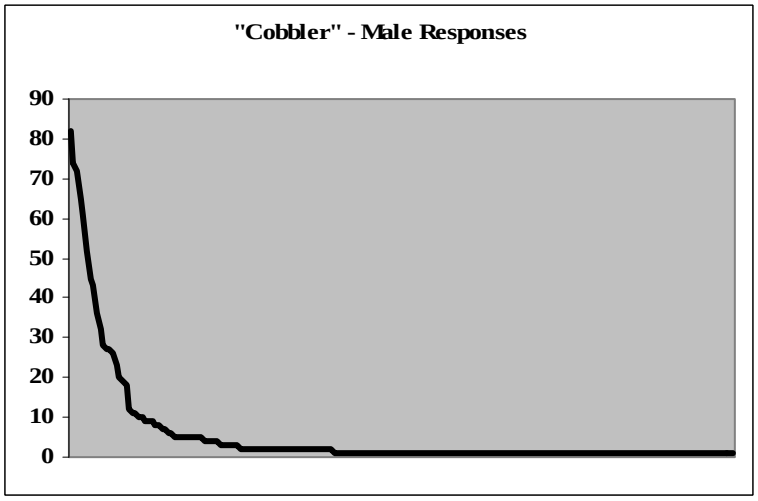

"Swamp" - NY/NJ/PA Responses
450
400
350
300
250
200
150
100
50
0

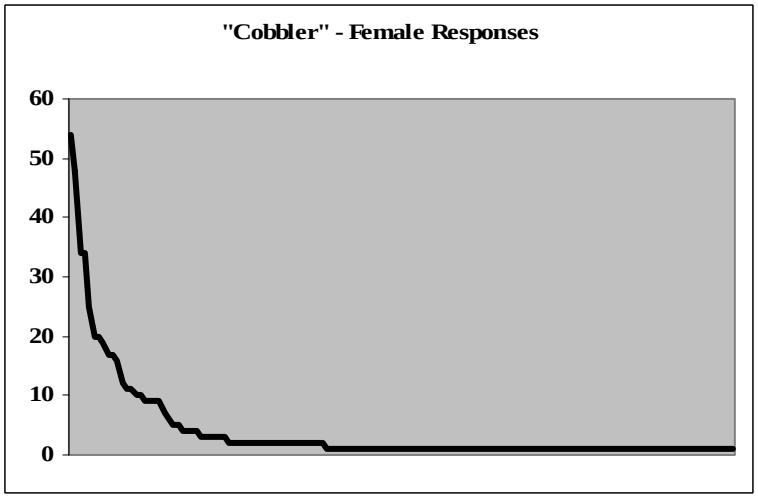

"Swamp" - MD/DC/DE/VA/WV Responses
400
350
300
250
200
150
100
50
0

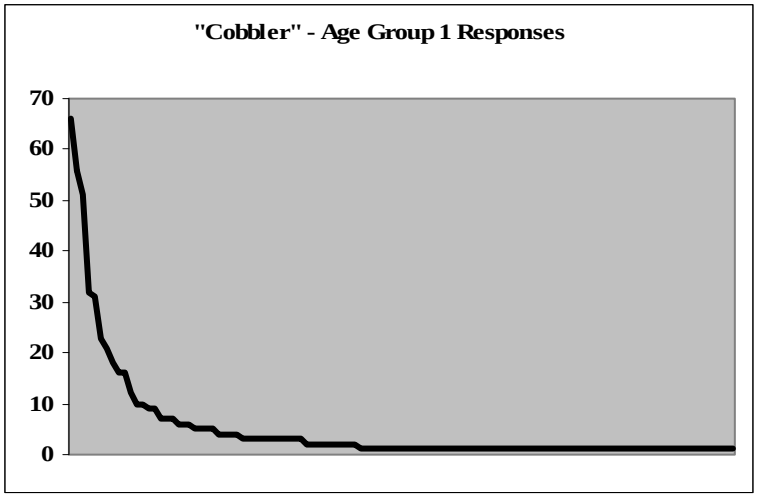

"Swamp" - NC/SC/GA/FL Responses
500
450
400
350
300
250
200
150
100
50
0

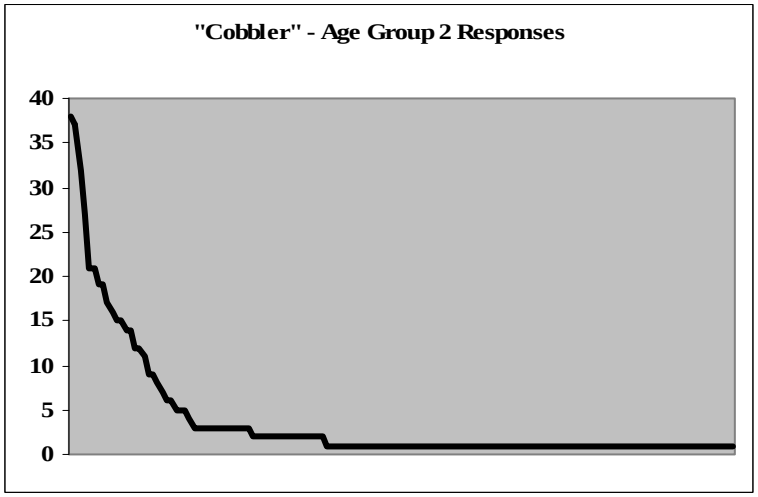
“Cobbler” Frequency Graphs

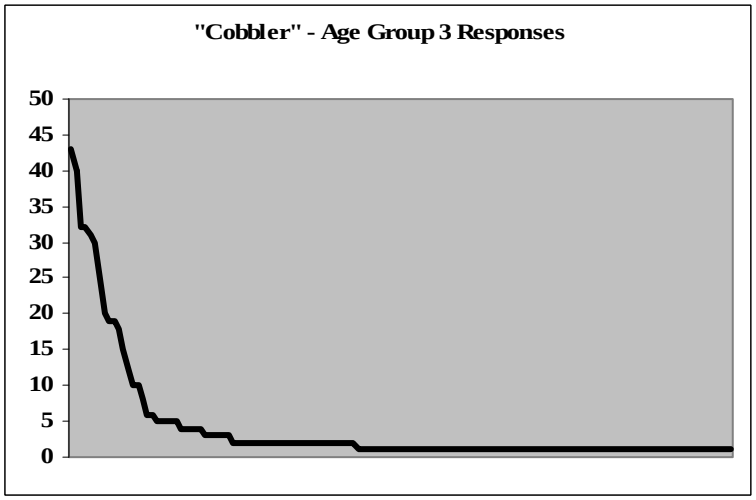


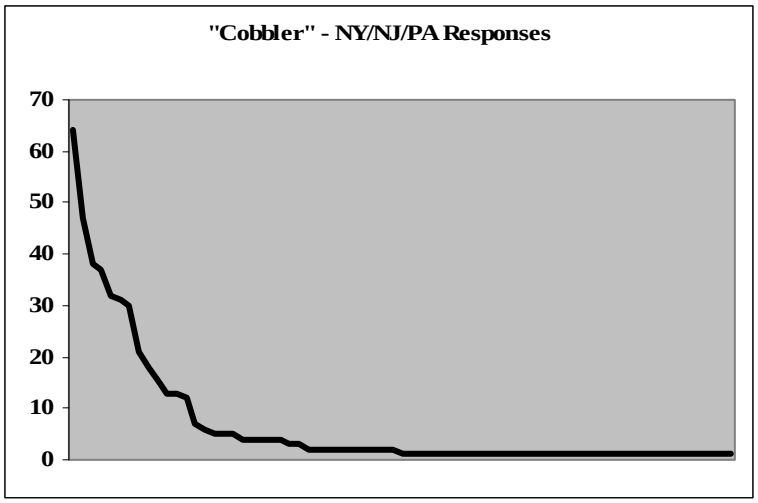


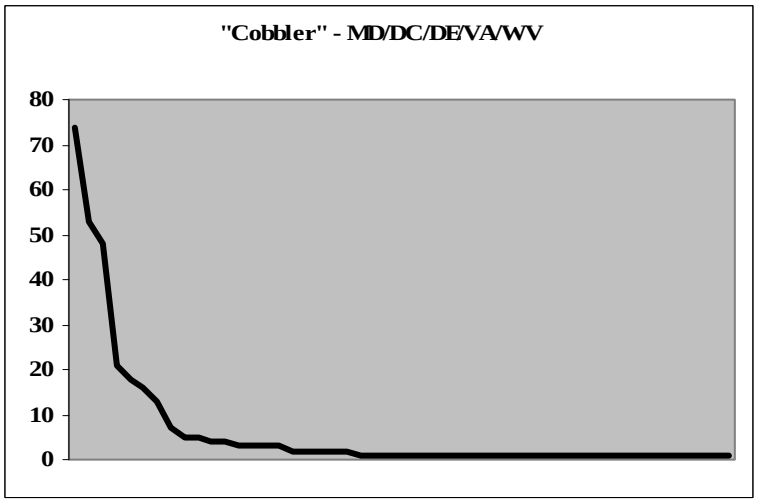


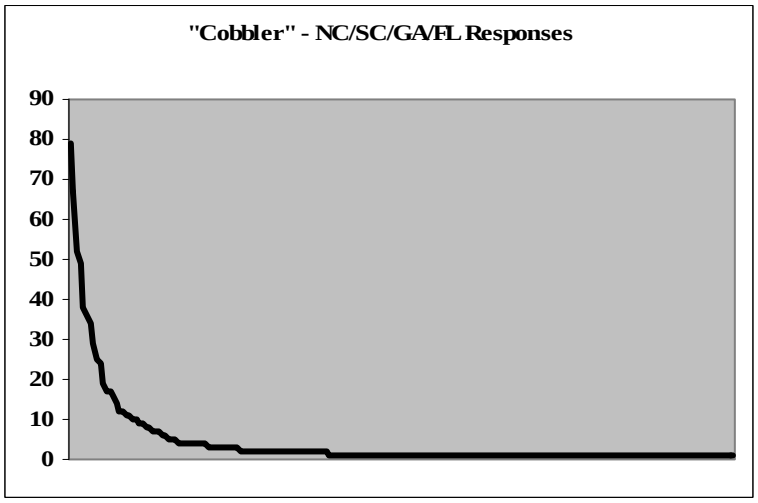

"Cobbler" - Age Group 1 Responses
70
60
50
40
30
20
10
0


"Cobbler" - Age Group 2 Responses
40
35
30
25
20
15
10
5
0

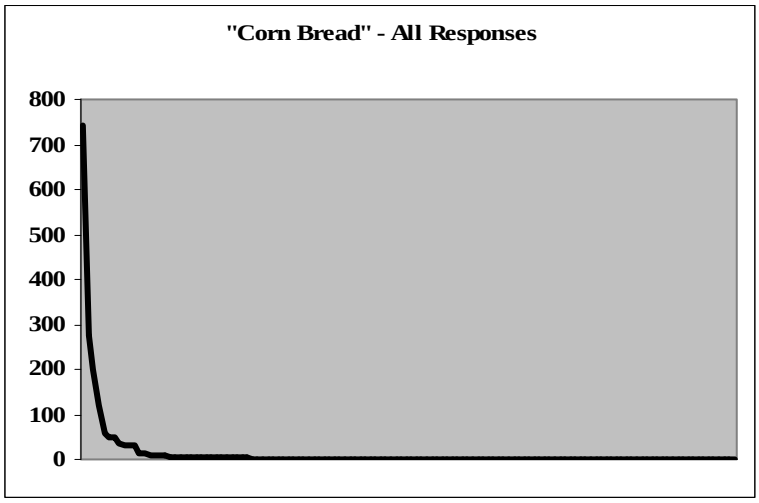

"Cobbler" - Age Group 3 Responses
50
45
40
35
30
25
20
15
10
5
0

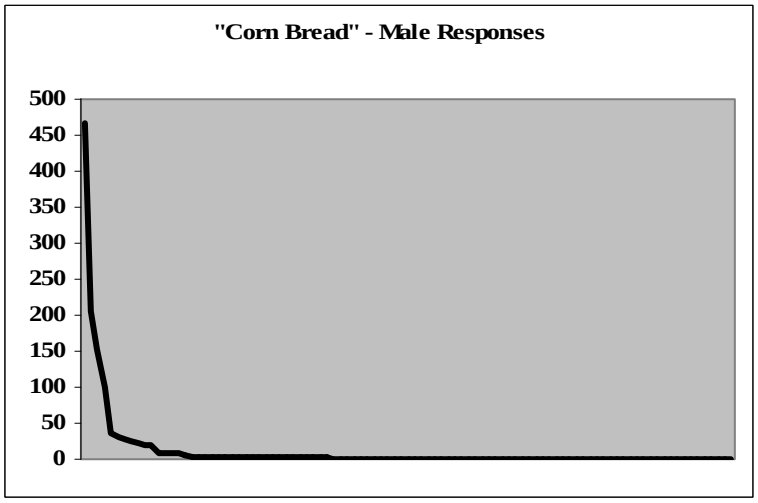

"Cobbler" - NY/NJ/PA Responses
70
60
50
40
30
20
10
0

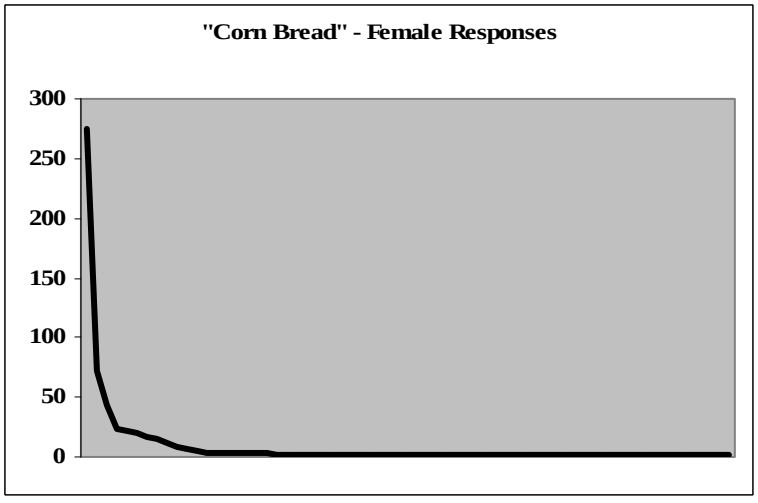

"Cobbler" - MD/DC/DE/VA/WV
80
70
60
50
40
30
20
10
0

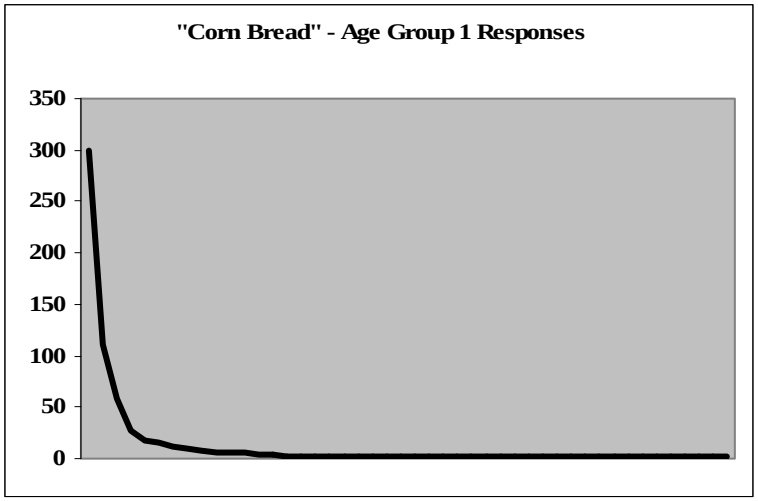

"Cobbler" - NC/SC/GA/FL Responses
90
80
70
60
50
40
30
20
10
0

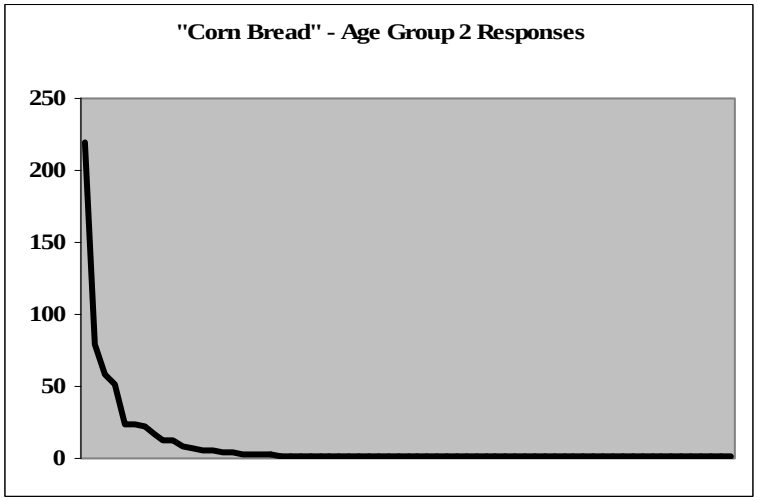
“Corn Bread” Frequency Graphs

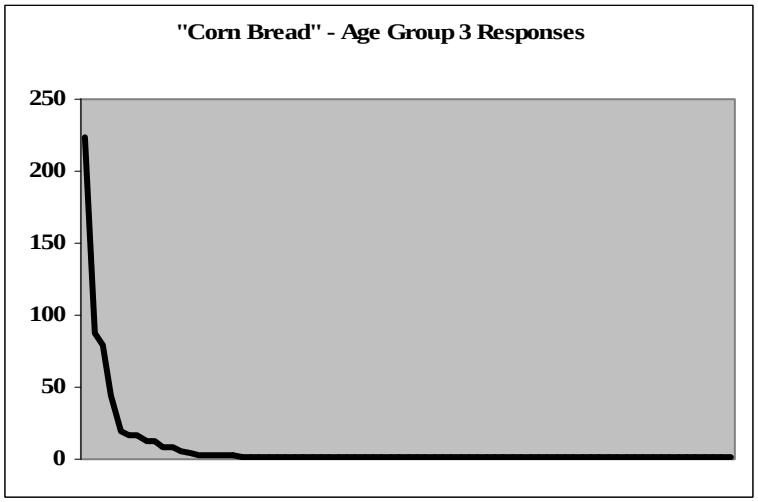


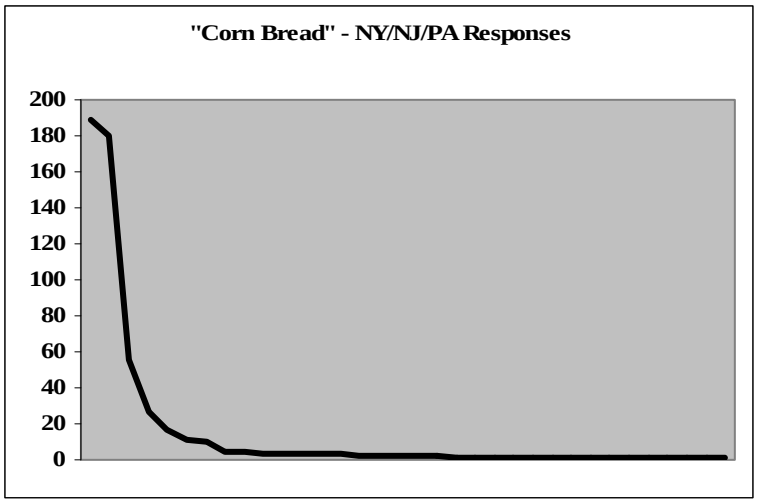


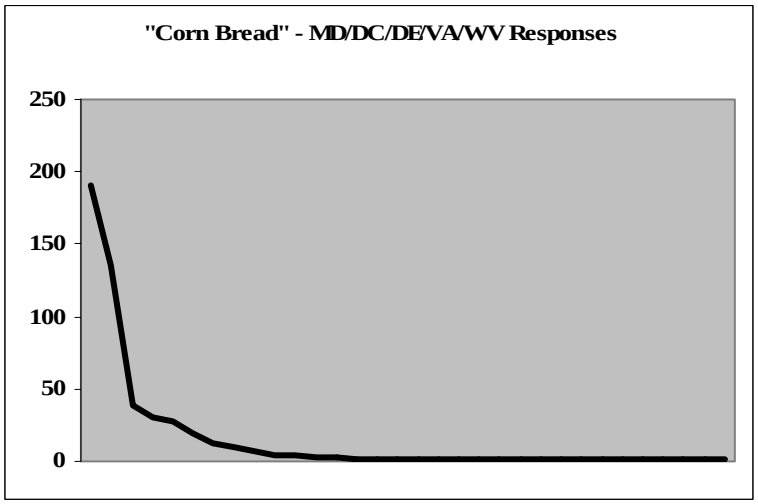


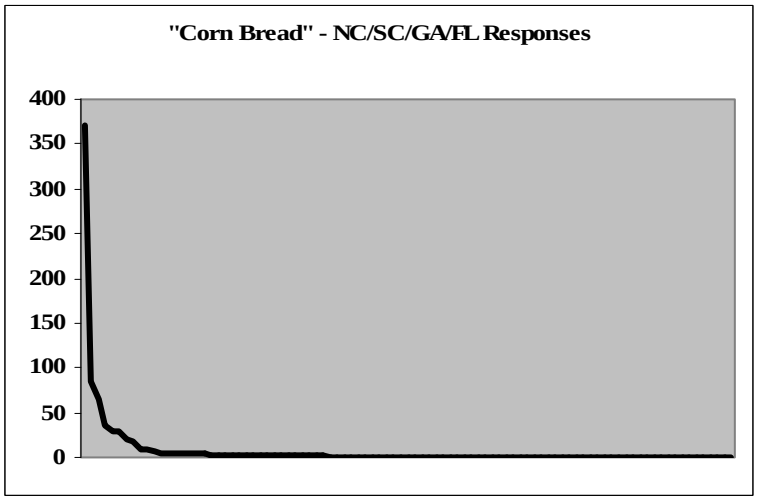

"Corn Bread" - Age Group 1 Responses
350
300
250
200
150
100
50
0


"Corn Bread" - Age Group 2 Responses
250
200
150
100
50
0

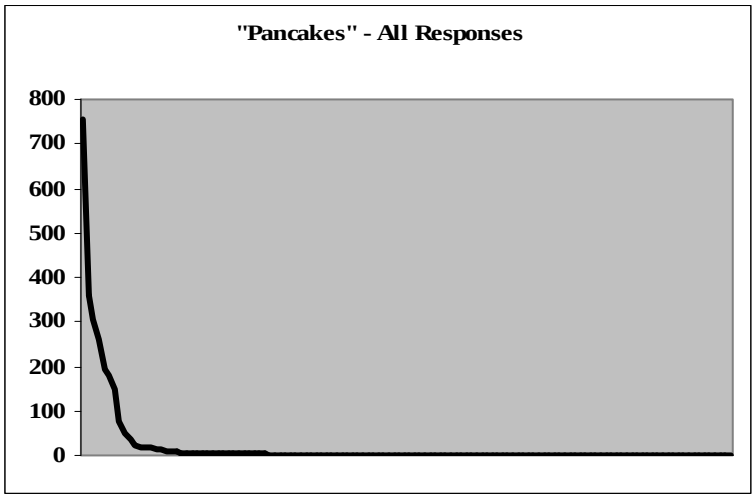

"Corn Bread" - Age Group 3 Responses
250
200
150
100
50
0

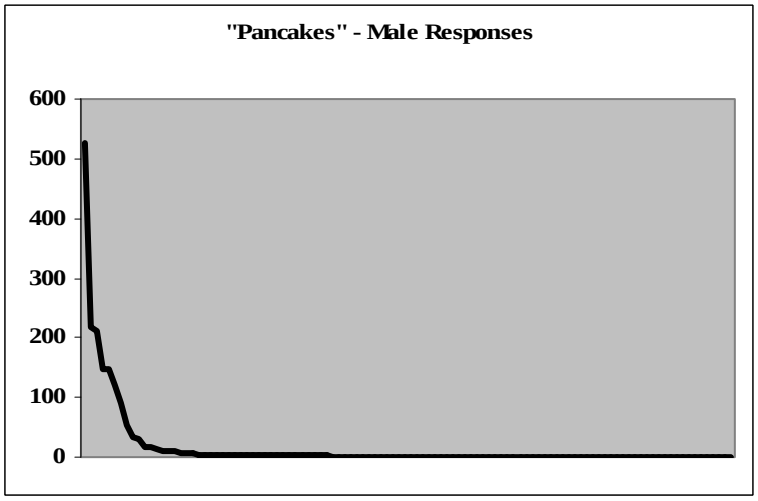

"Corn Bread" - NY/NJ/PA Responses
200
180
160
140
120
100
80
60
40
20
0

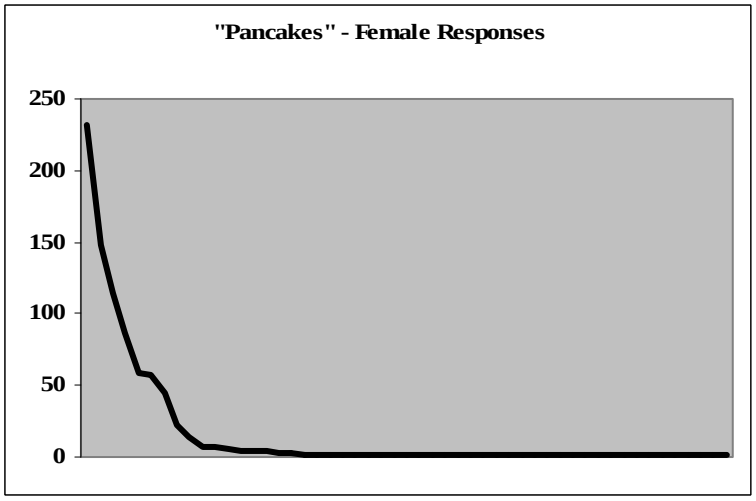

"Corn Bread" - MD/DC/DE/VA/WV Responses
250
200
150
100
50
0

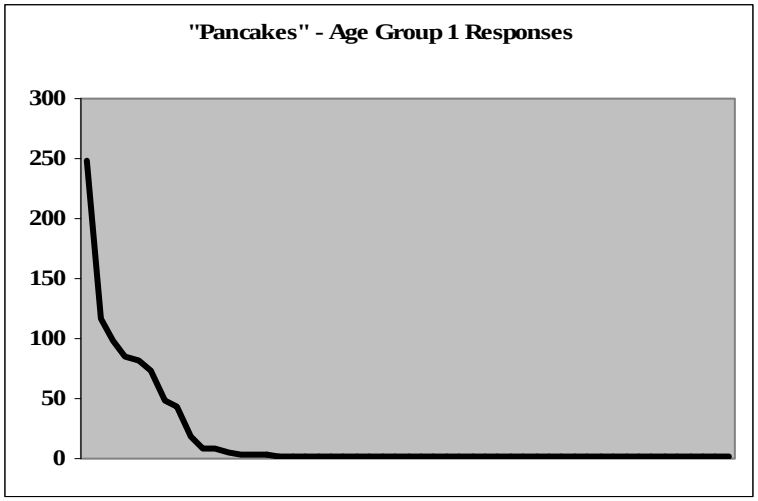

"Corn Bread" - NC/SC/GA/FL Responses
400
350
300
250
200
150
100
50
0

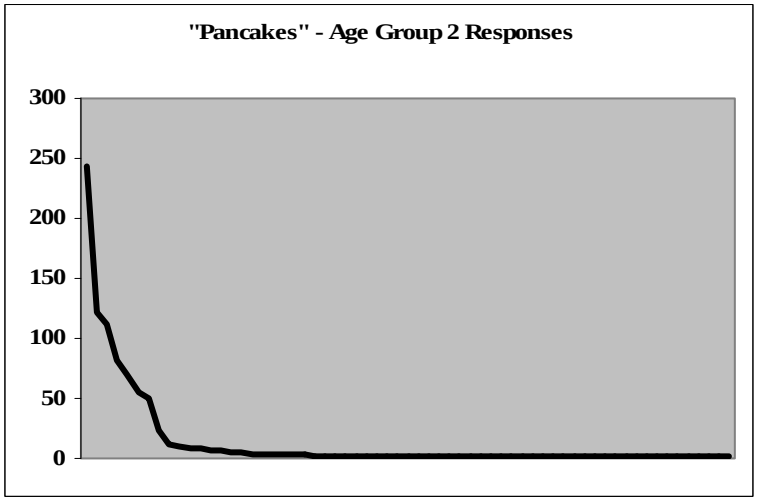
“Pancakes” Frequency Graphs

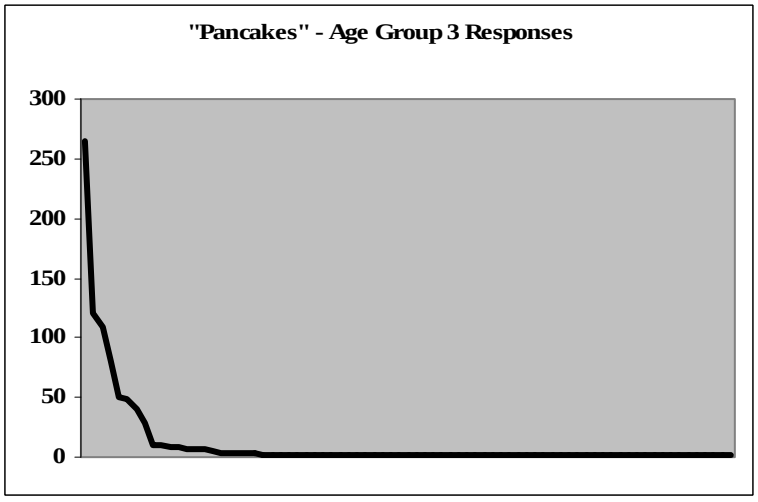


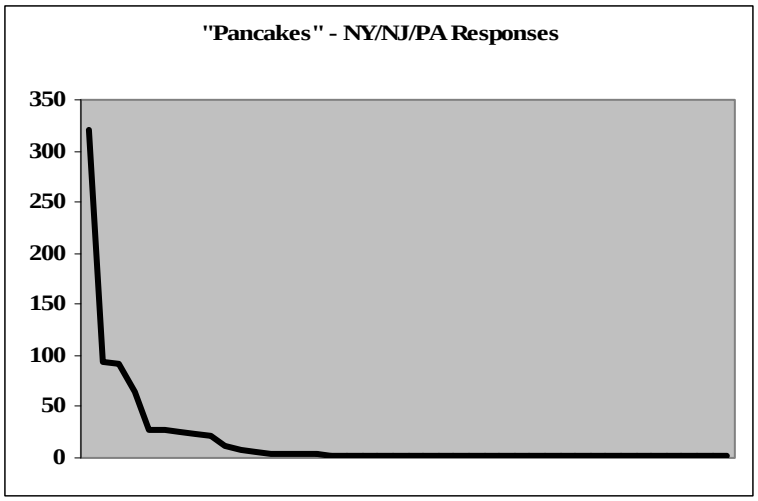


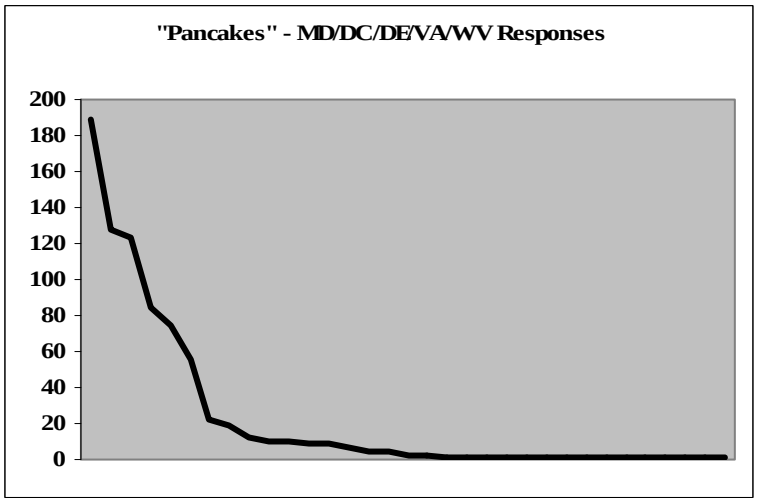


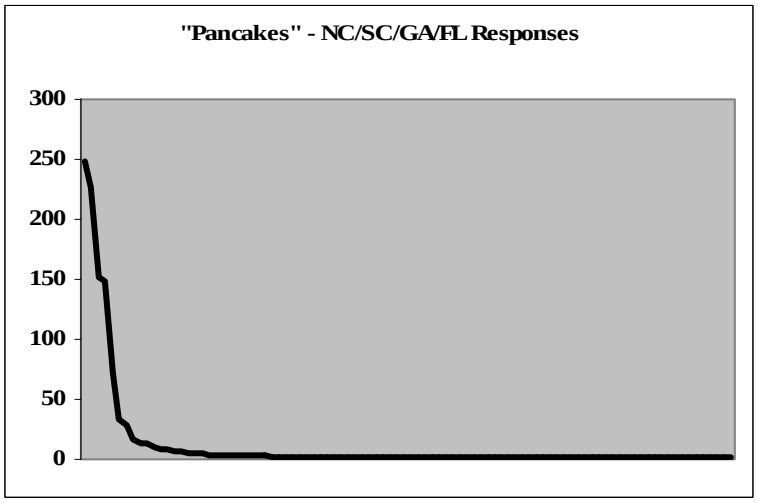

"Pancakes" - Age Group 1 Responses
300
250
200
150
100
50
0


"Pancakes" - Age Group 2 Responses
300
250
200
150
100
50
0

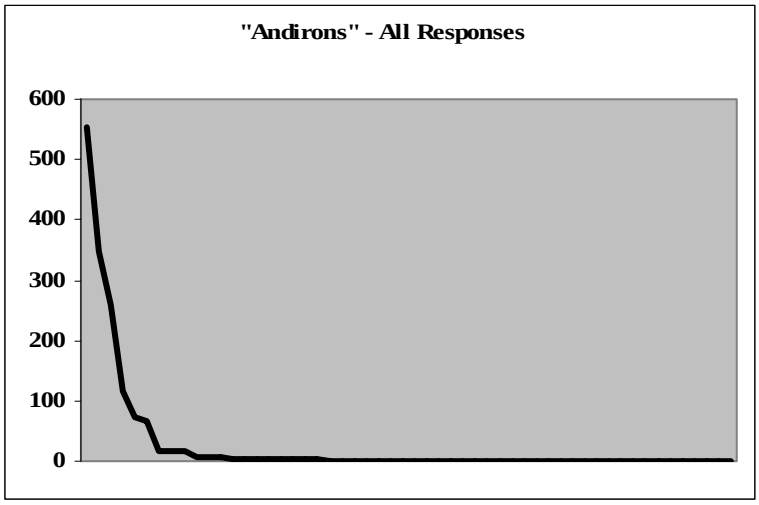

"Pancakes" - Age Group 3 Responses
300
250
200
150
100
50
0

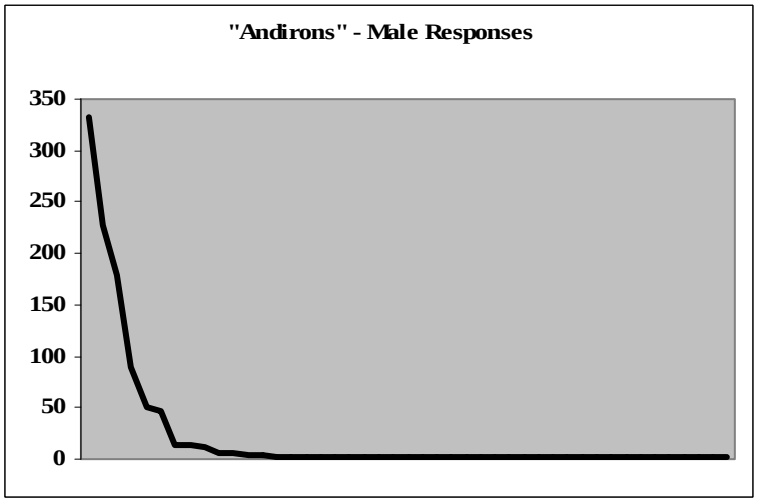

"Pancakes" - NY/NJ/PA Responses
350
300
250
200
150
100
50
0

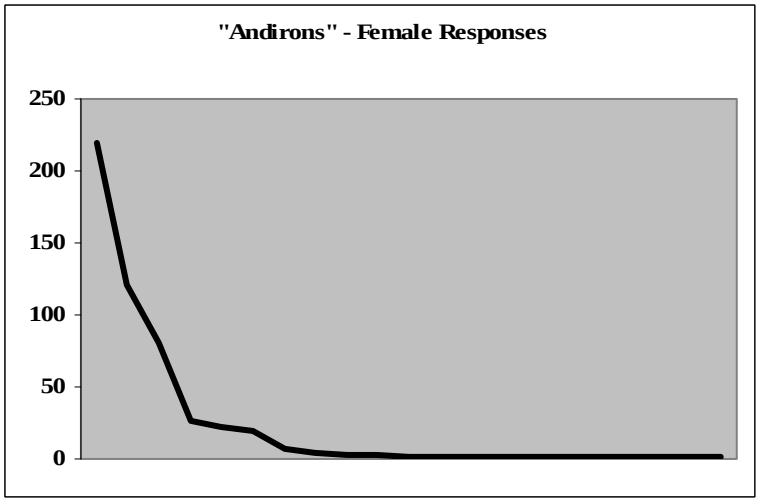

"Pancakes" - MD/DC/DE/VA/WV Responses
200
180
160
140
120
100
80
60
40
20
0

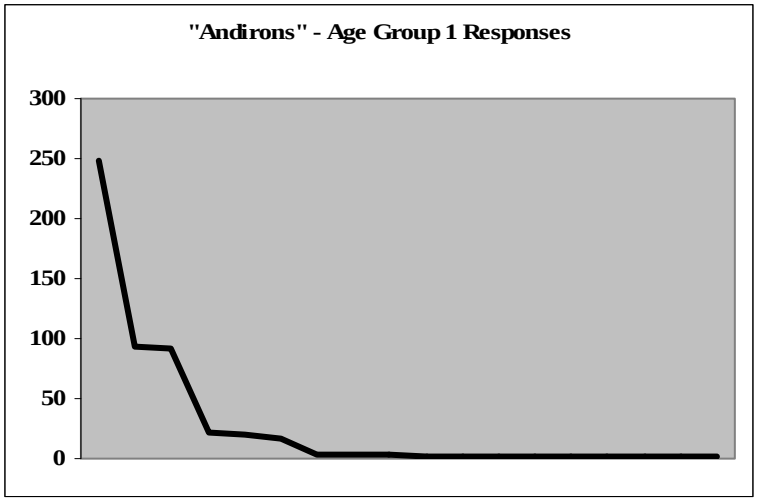

"Pancakes" - NC/SC/GA/FL Responses
300
250
200
150
100
50
0

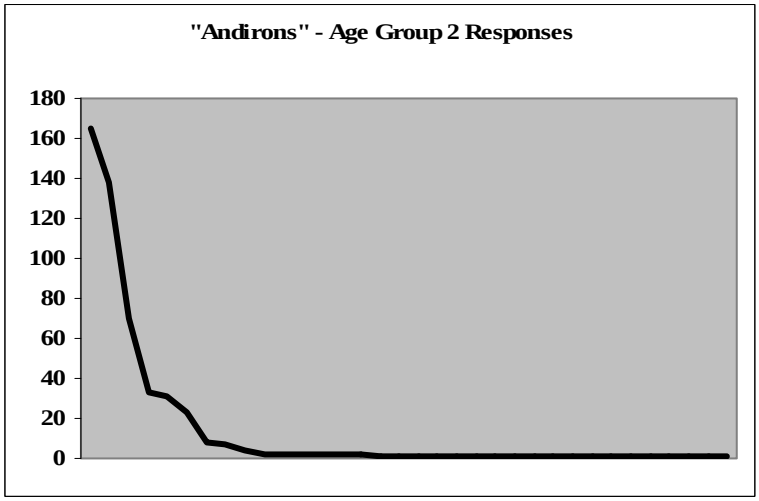
“Andirons” Frequency Graphs

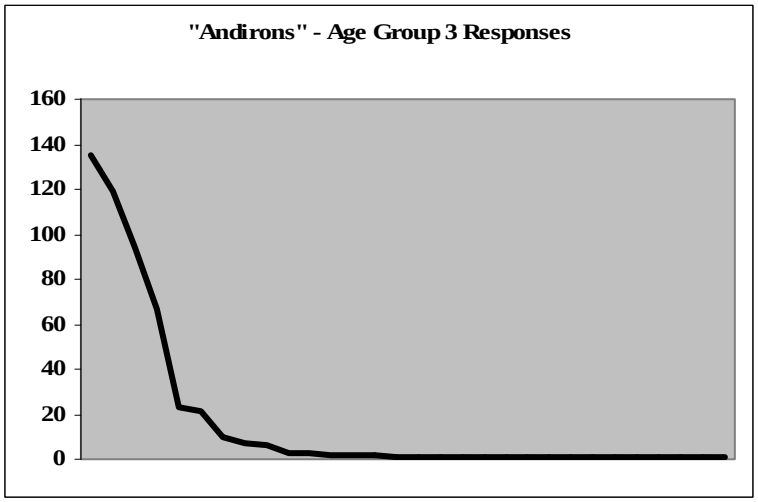


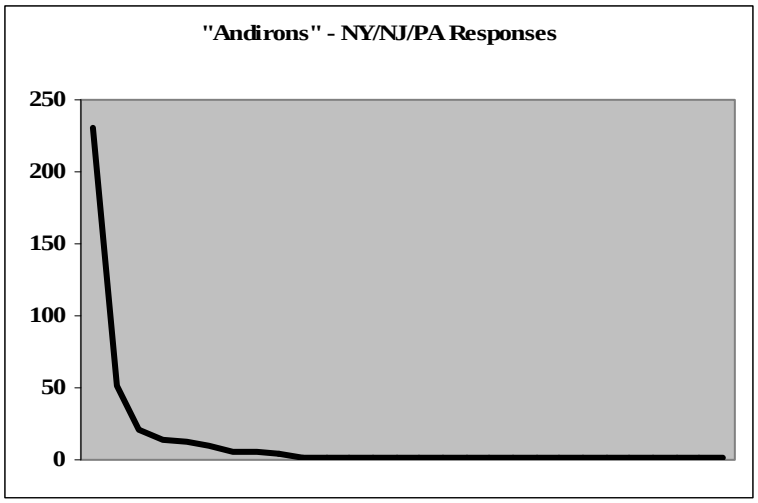


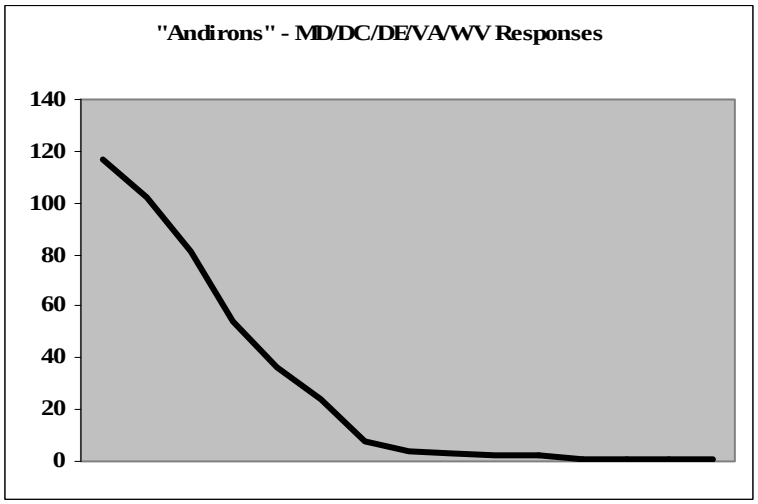


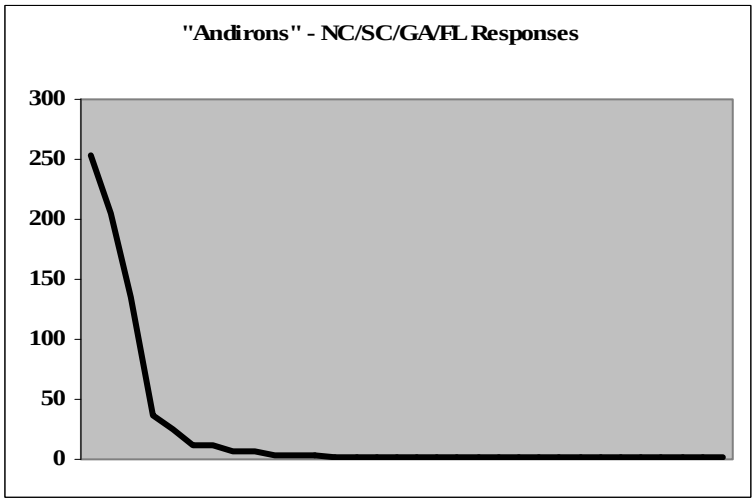

"Andirons" - Age Group 1 Responses
300
250
200
150
100
50
0


"Andirons" - Age Group 2 Responses
180
160
140
120
100
80
60
40
20
0

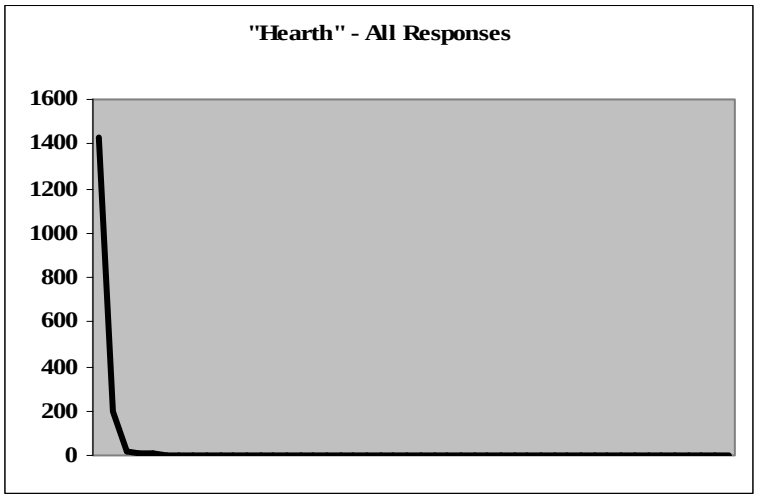

"Andirons" - Age Group 3 Responses
160
140
120
100
80
60
40
20
0

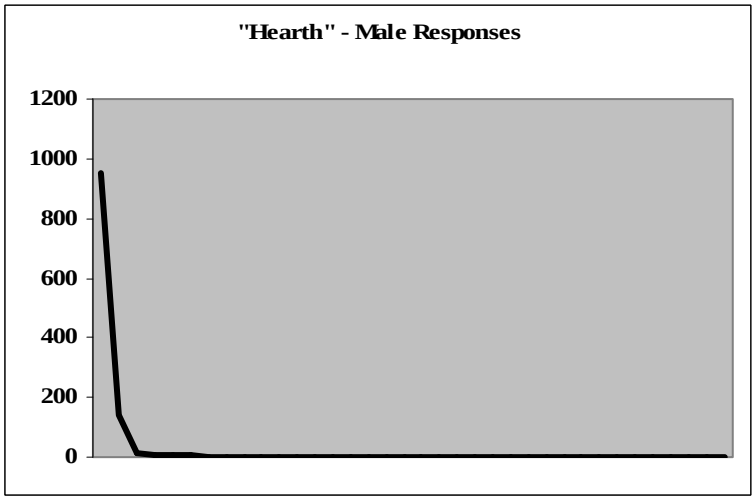

"Andirons" - NY/NJ/PA Responses
250
200
150
100
50
0

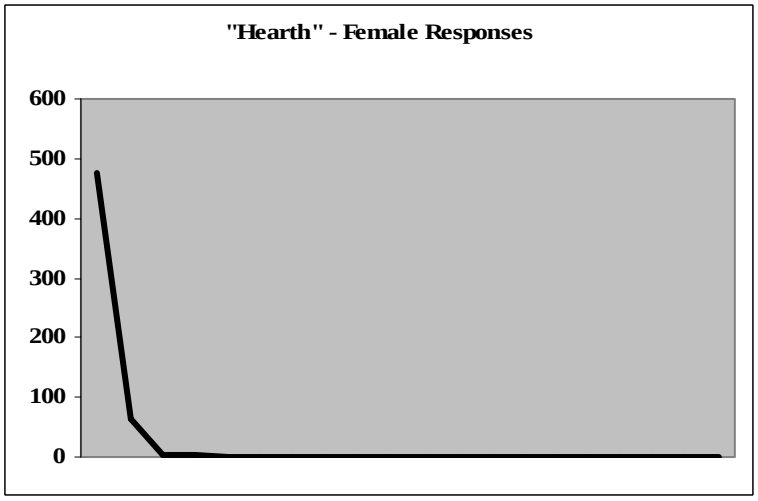

"Andirons" - MD/DC/DE/VA/WV Responses
140
120
100
80
60
40
20
0

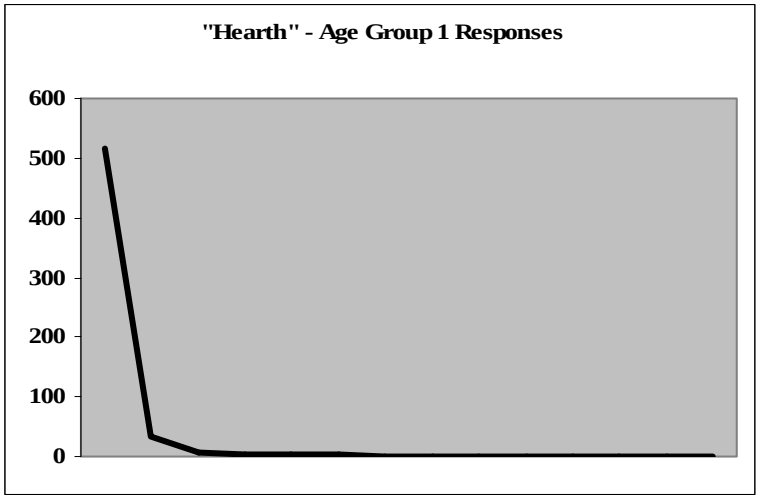

"Andirons" - NC/SC/GA/FL Responses
300
250
200
150
100
50
0

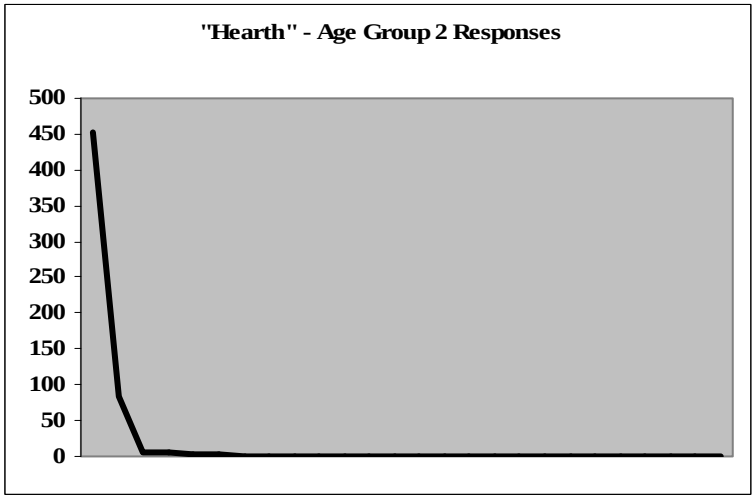
“Hearth” Frequency Graphs

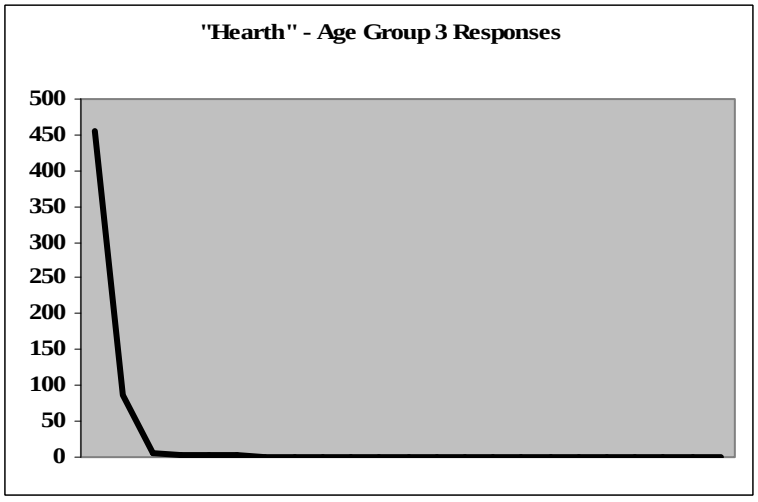


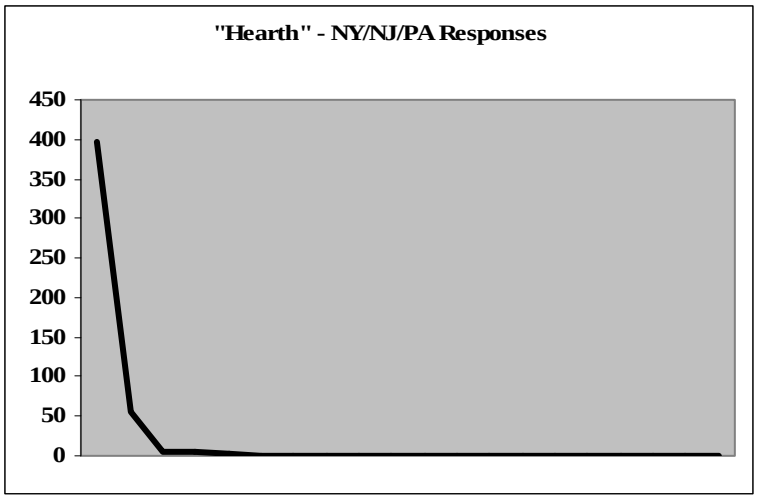


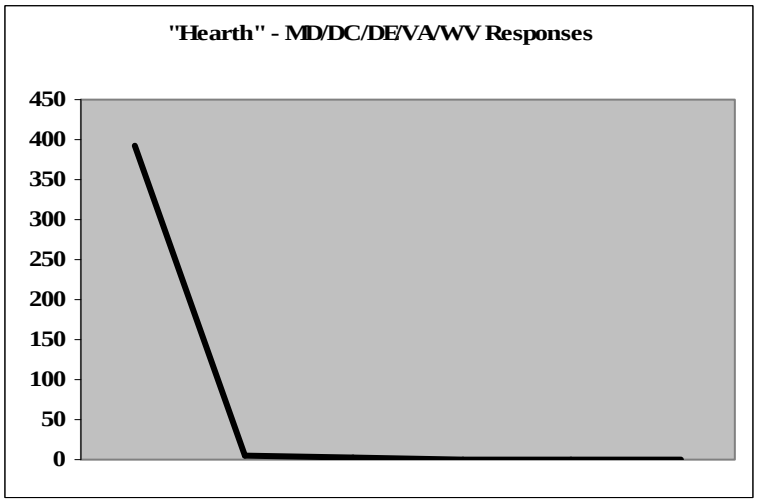


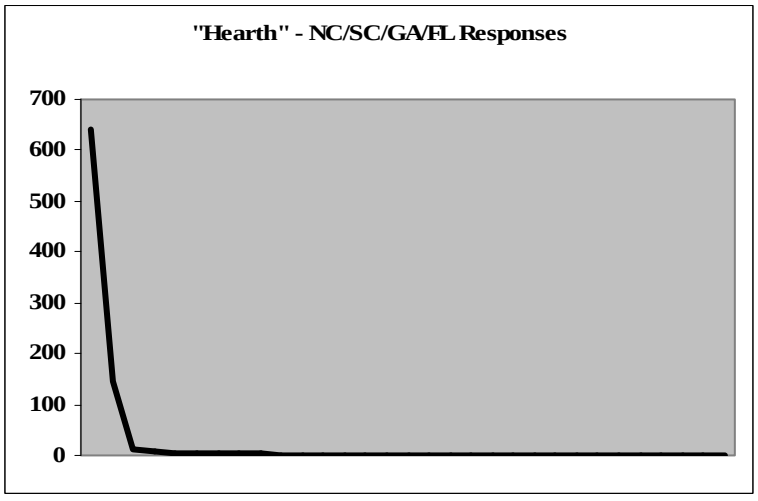

"Hearth" - Age Group 1 Responses
600
500
400
300
200
100
0


"Hearth" - Age Group 2 Responses
500
450
400
350
300
250
200
150
100
50
0

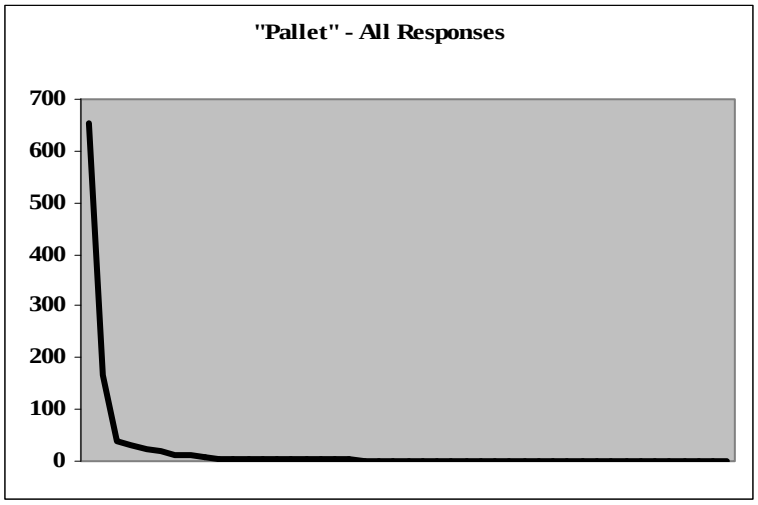

"Hearth" - Age Group 3 Responses
500
450
400
350
300
250
200
150
100
50
0

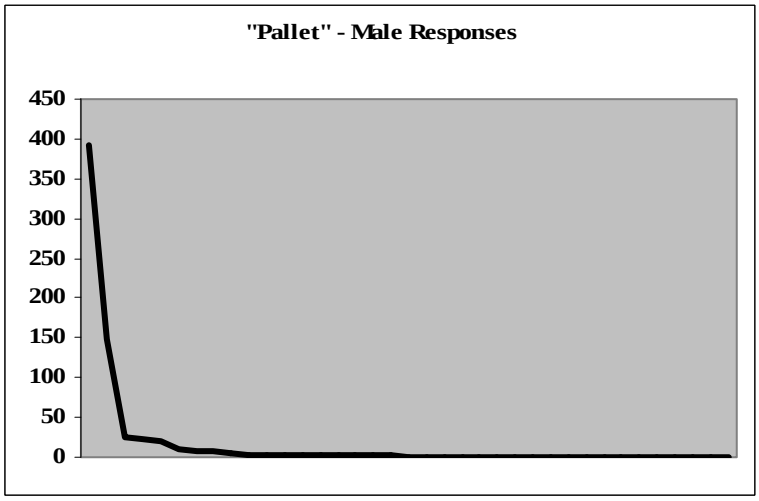

"Hearth" - NY/NJ/PA Responses
450
400
350
300
250
200
150
100
50
0

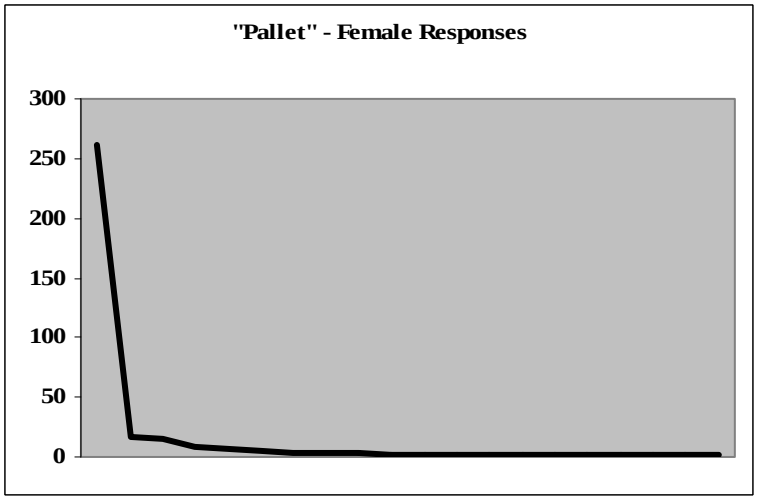

"Hearth" - MD/DC/DE/VA/WV Responses
450
400
350
300
250
200
150
100
50
0

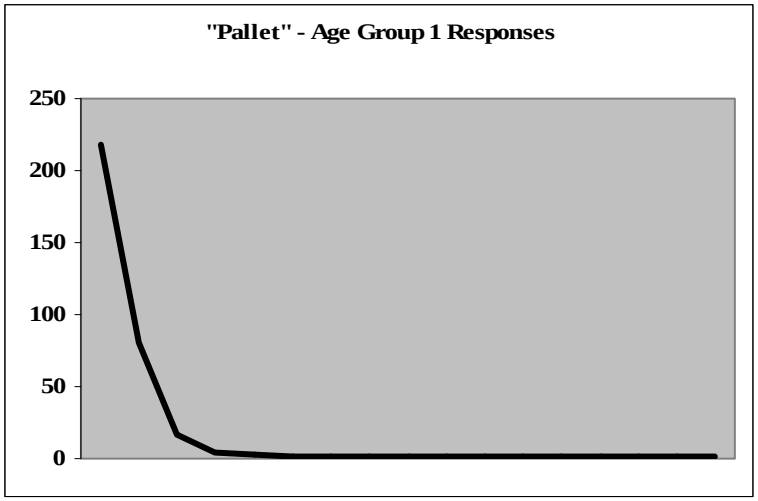

"Hearth" - NC/SC/GA/FL Responses
700
600
500
400
300
200
100
0

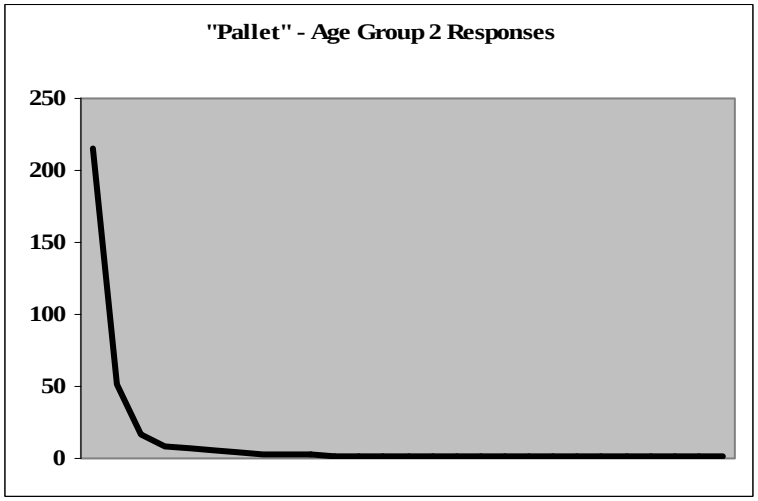
“Pallet” Frequency Graphs

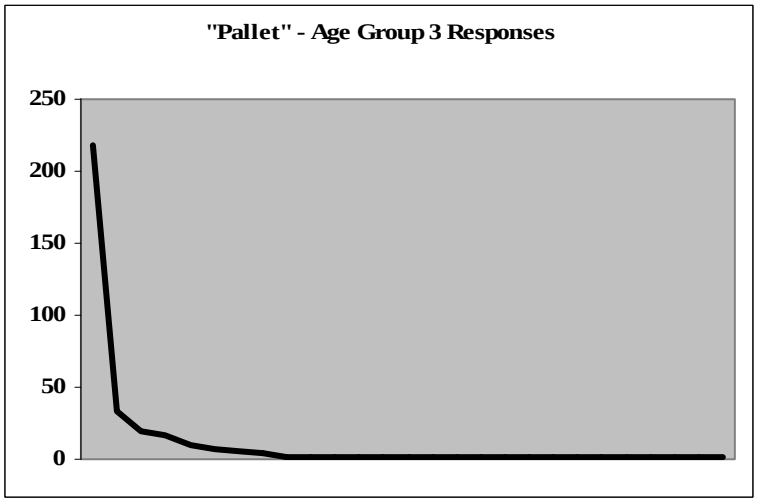


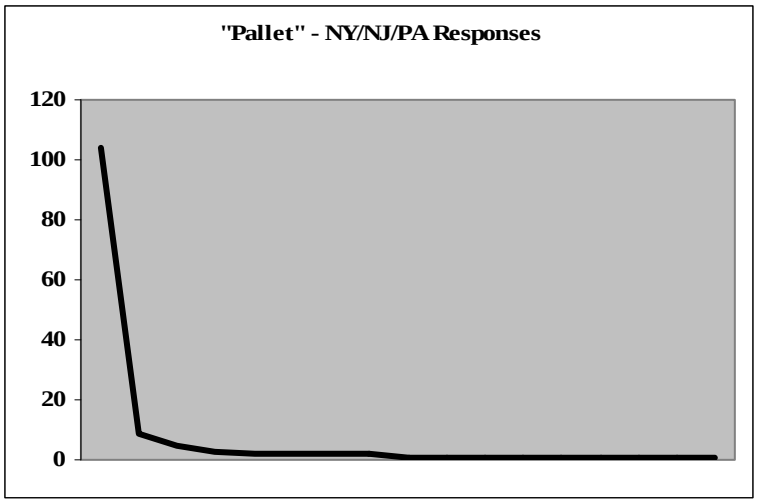


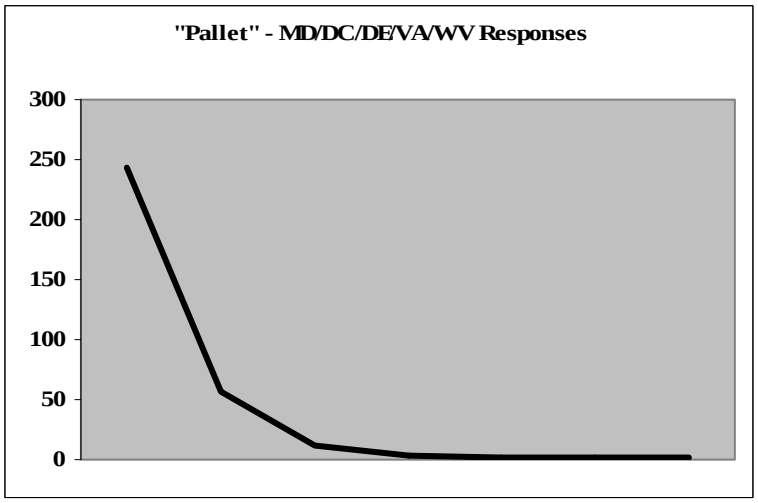


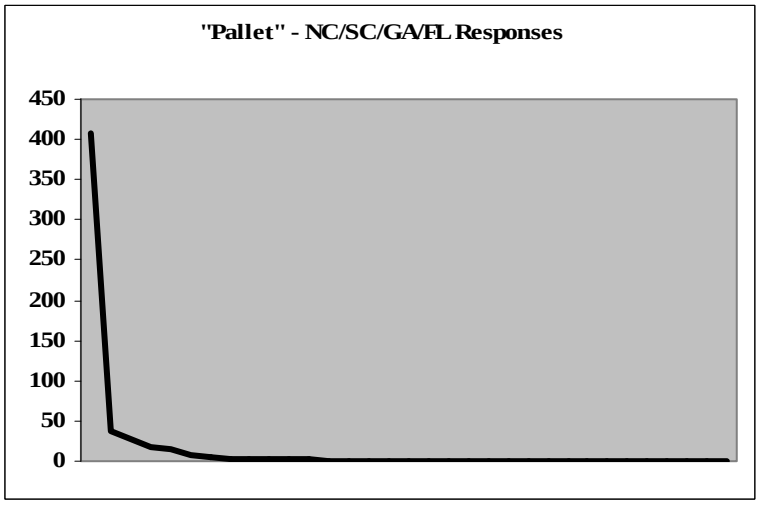

"Pallet" - Age Group 1 Responses
250
200
150
100
50
0


"Pallet" - Age Group 2 Responses
250
200
150
100
50
0

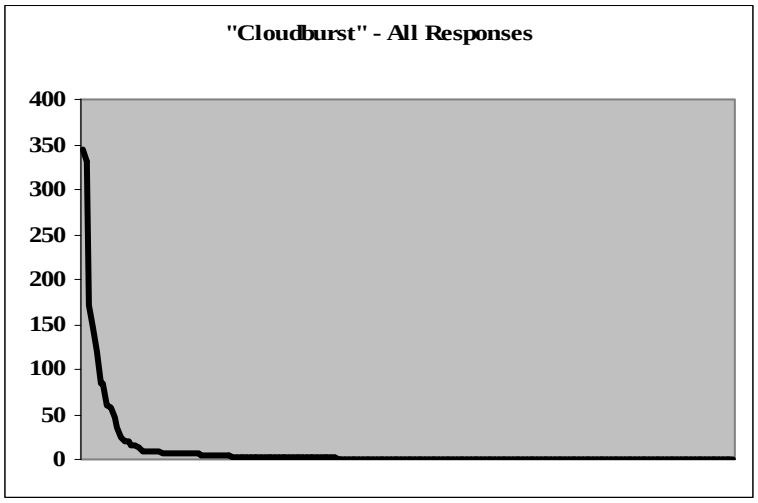

"Pallet" - Age Group 3 Responses
250
200
150
100
50
0

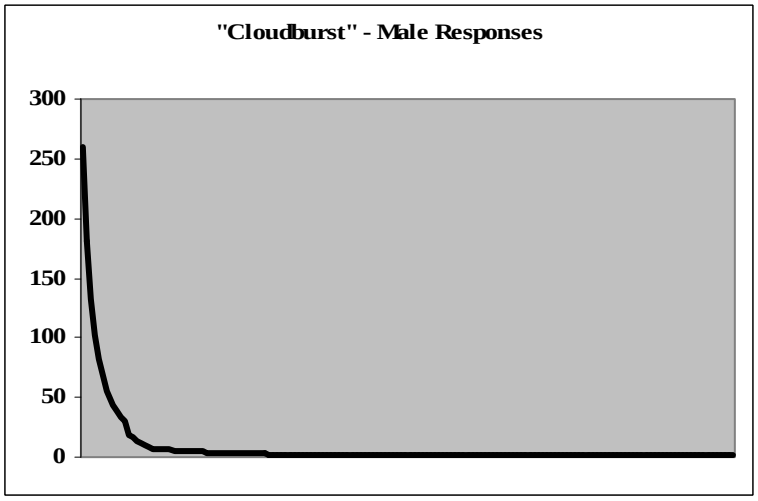

"Pallet" - NY/NJ/PA Responses
120
100
80
60
40
20
0

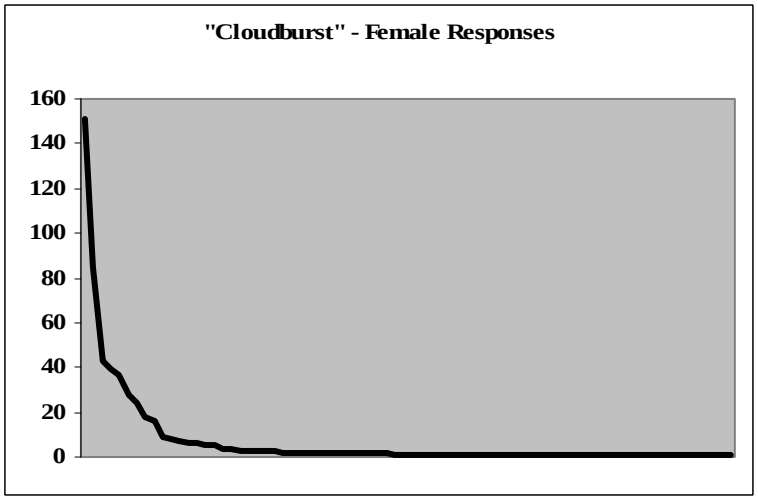

"Pallet" - MD/DC/DE/VA/WV Responses
300
250
200
150
100
50
0

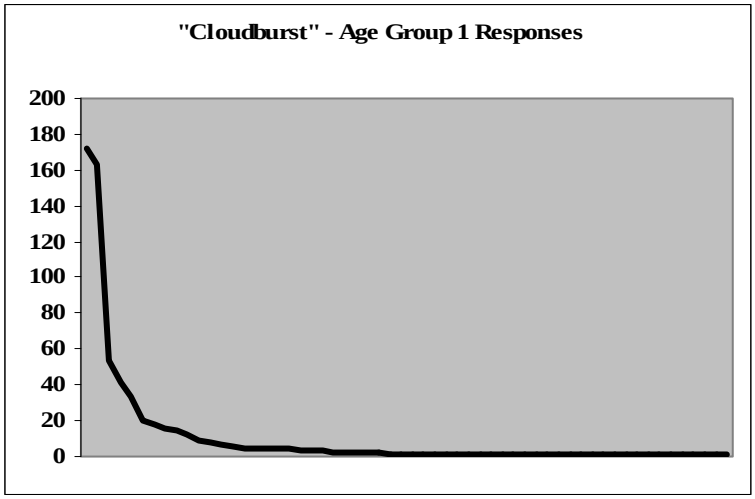

"Pallet" - NC/SC/GA/FL Responses
450
400
350
300
250
200
150
100
50
0

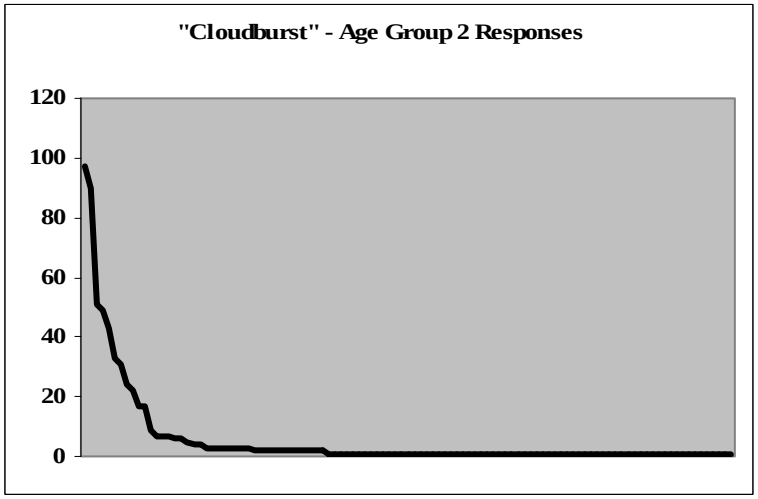
“Cloudburst” Frequency Graphs

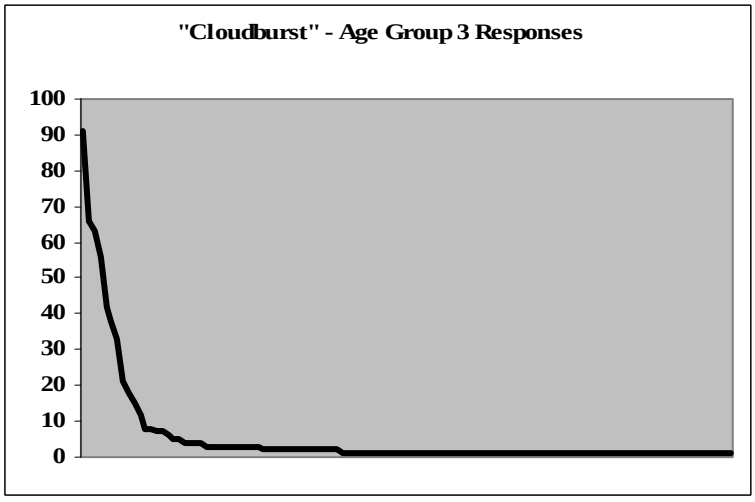


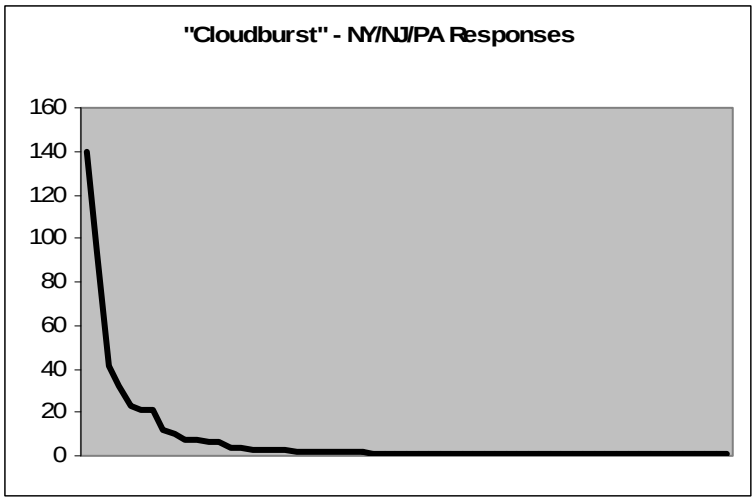


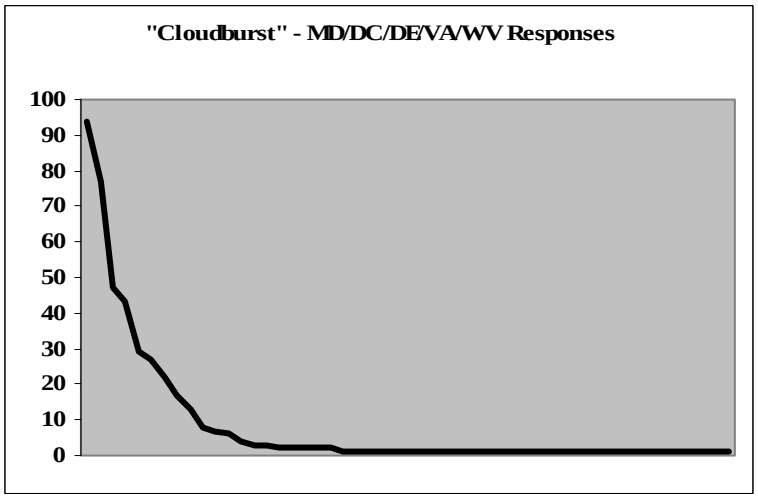


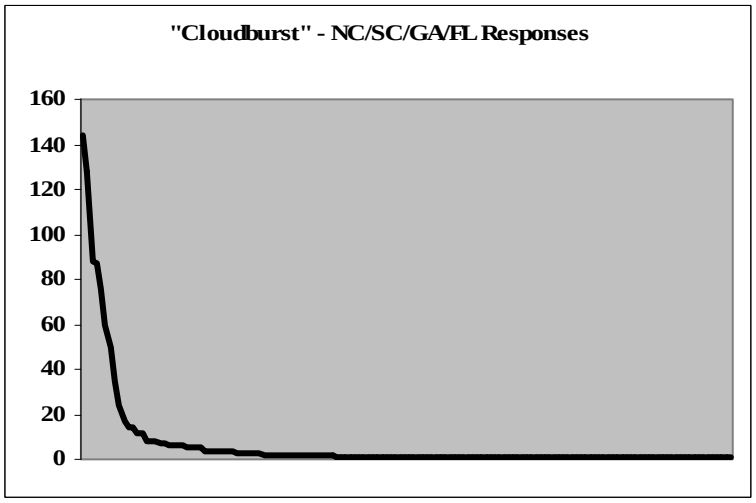

"Cloudburst" - Age Group 1 Responses
200
180
160
140
120
100
80
60
40
20
0


"Cloudburst" - Age Group 2 Responses
120
100
80
60
40
20
0

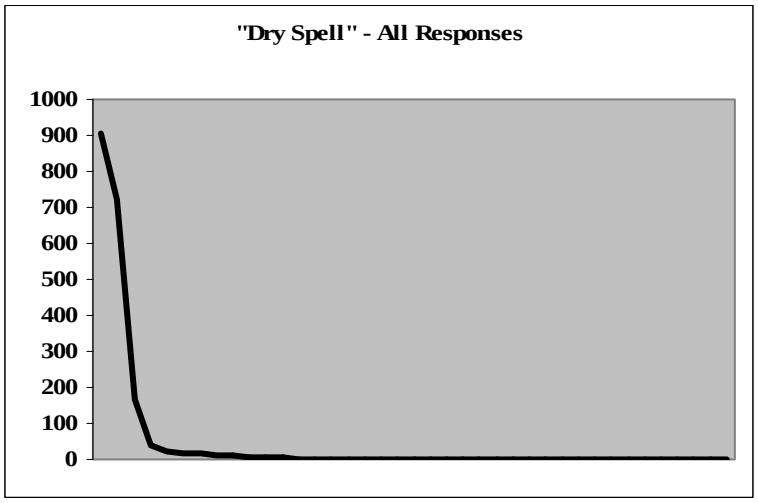

"Cloudburst" - Age Group 3 Responses
100
90
80
70
60
50
40
30
20
10
0

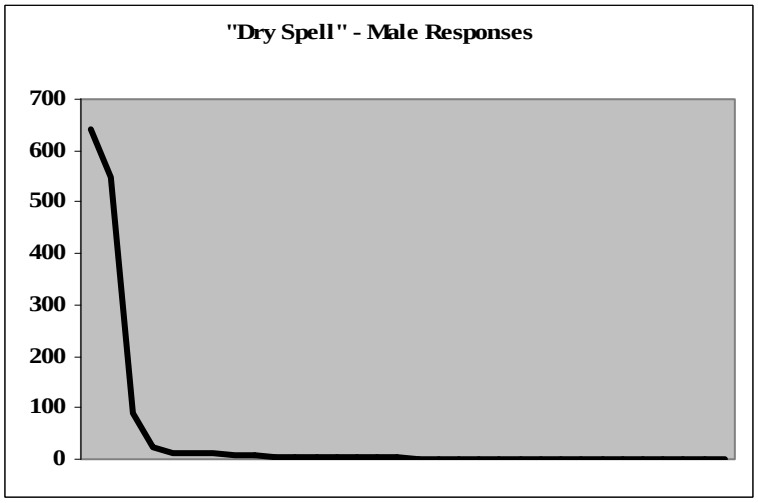

"Cloudburst" - NY/NJ/PA Responses
160
140
120
100
80
60
40
20
0

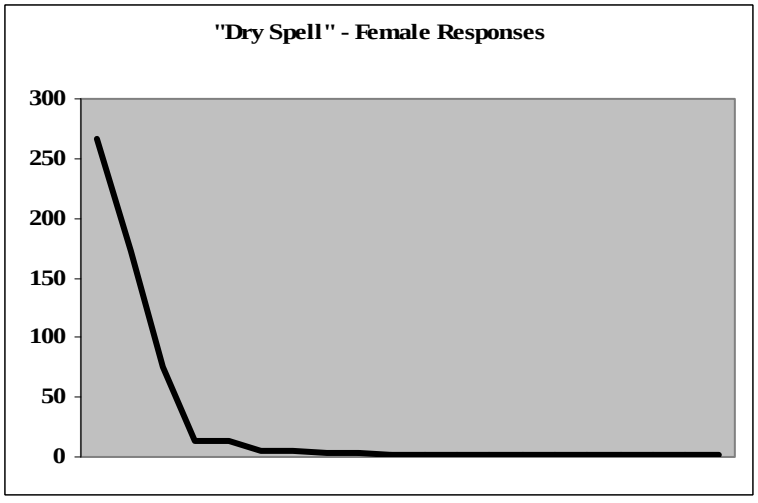

"Cloudburst" - MD/DC/DE/VA/WV Responses
100
90
80
70
60
50
40
30
20
10
0

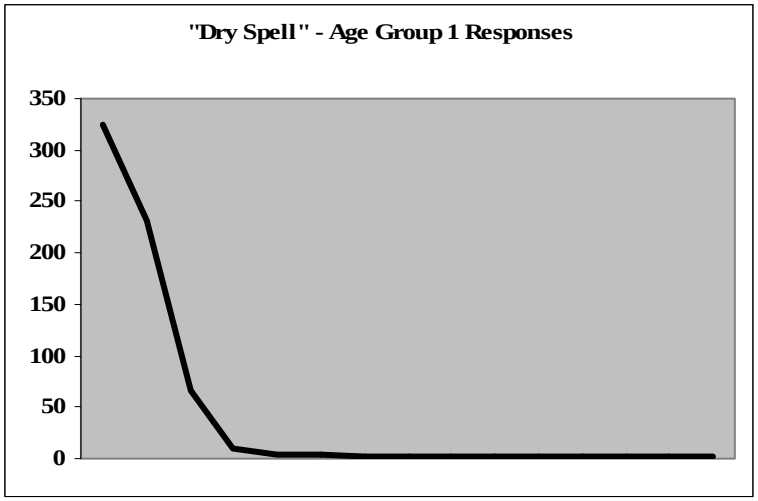

"Cloudburst" - NC/SC/GA/FL Responses
160
140
120
100
80
60
40
20
0

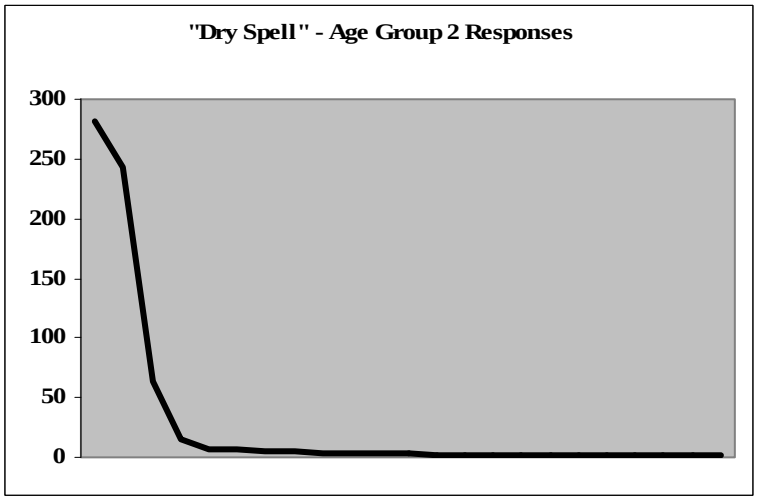
“Dry Spell” Frequency Graphs

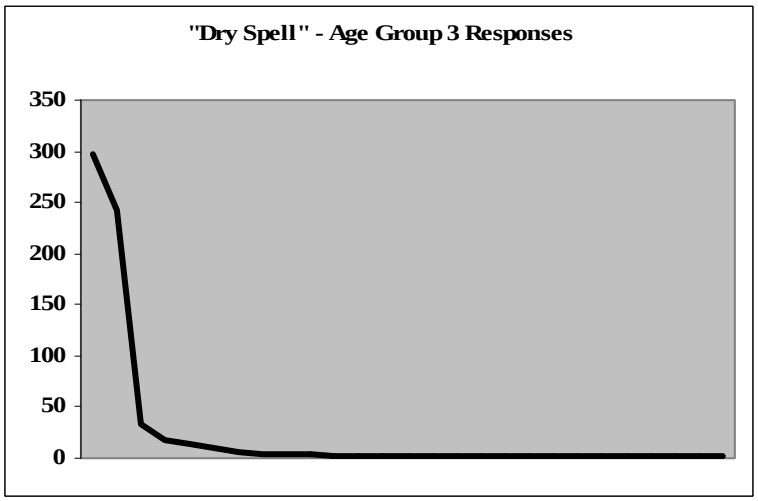


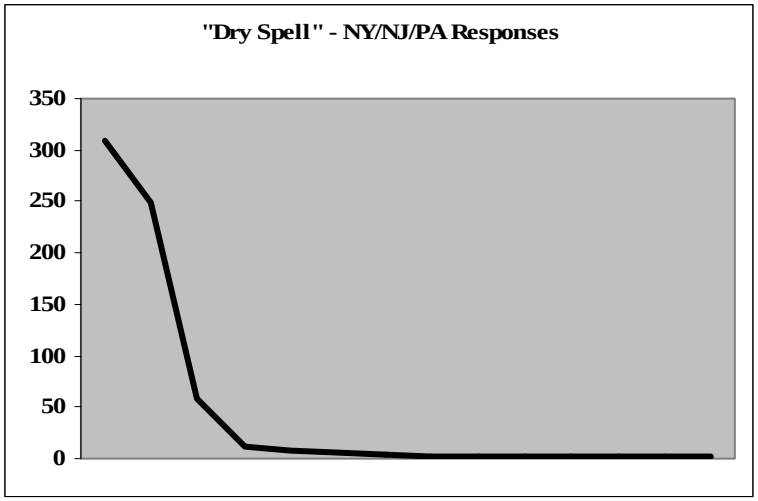


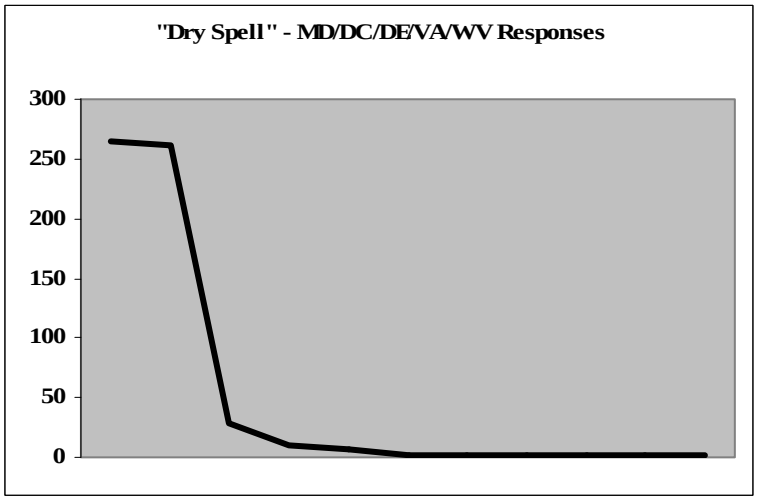


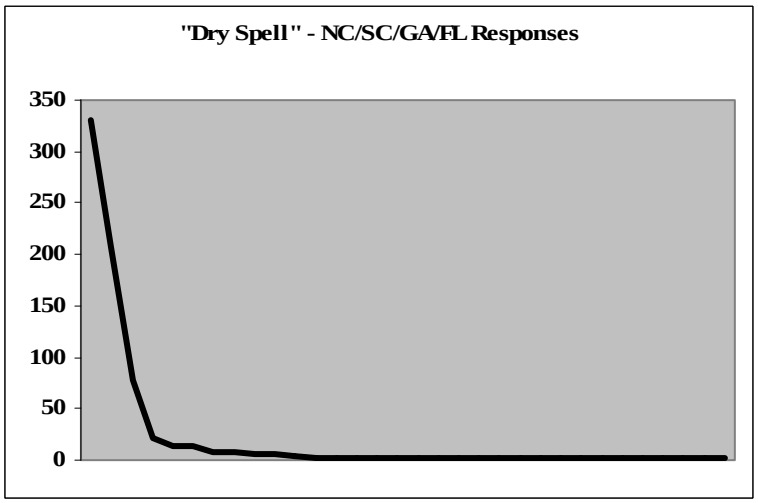

"Dry Spell" - Age Group 1 Responses
350
300
250
200
150
100
50
0


"Dry Spell" - Age Group 2 Responses
300
250
200
150
100
50
0

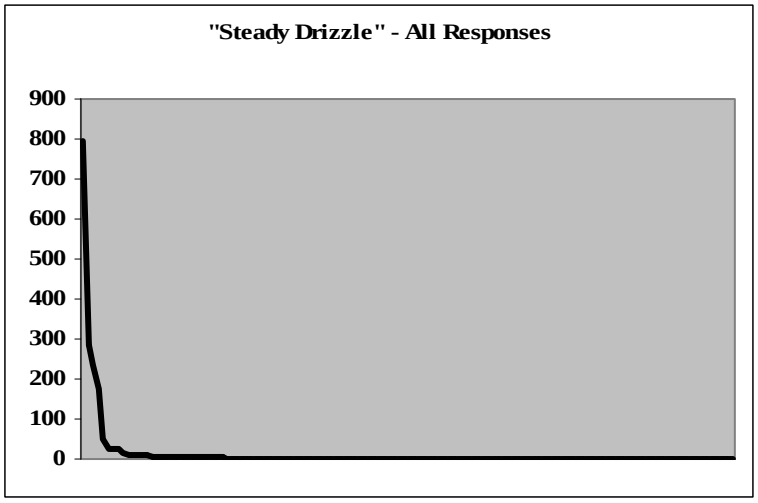

"Dry Spell" - Age Group 3 Responses
350
300
250
200
150
100
50
0

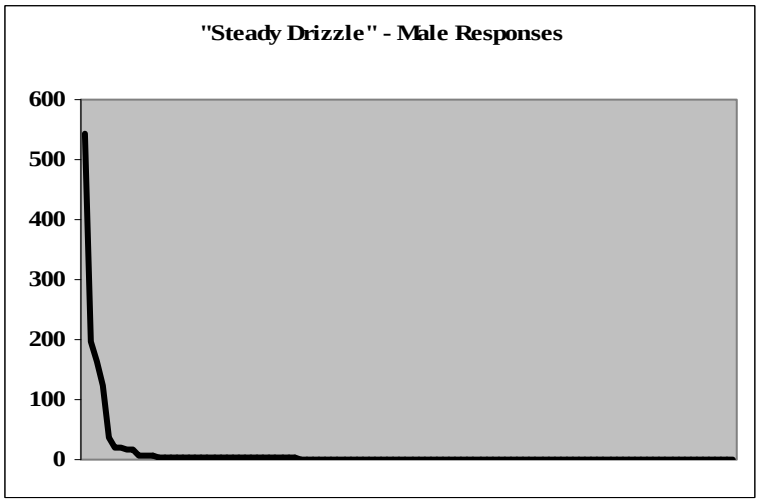

"Dry Spell" - NY/NJ/PA Responses
350
300
250
200
150
100
50
0

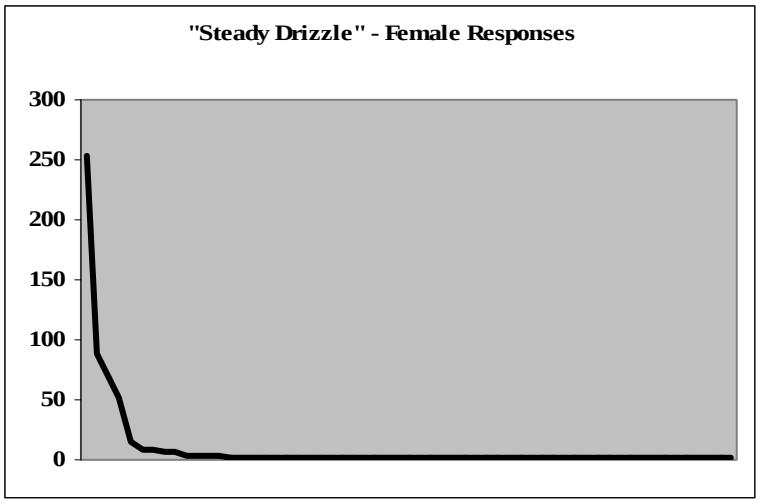

"Dry Spell" - MD/DC/DE/VA/WV Responses
300
250
200
150
100
50
0

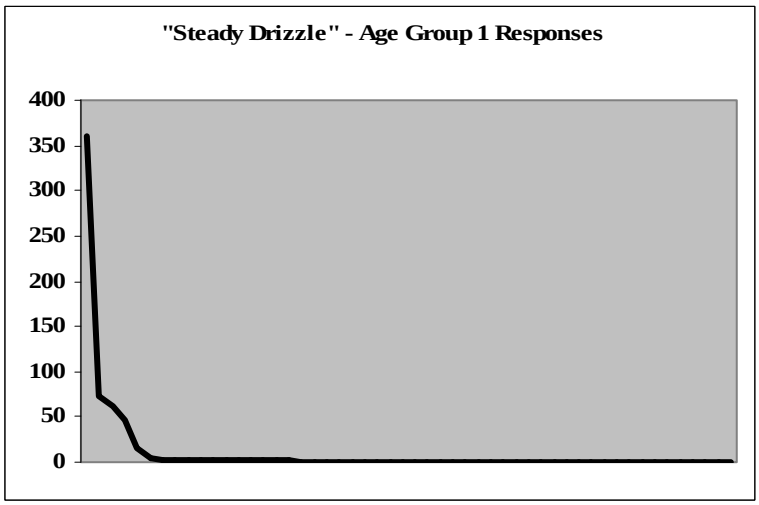

"Dry Spell" - NC/SC/GA/FL Responses
350
300
250
200
150
100
50
0

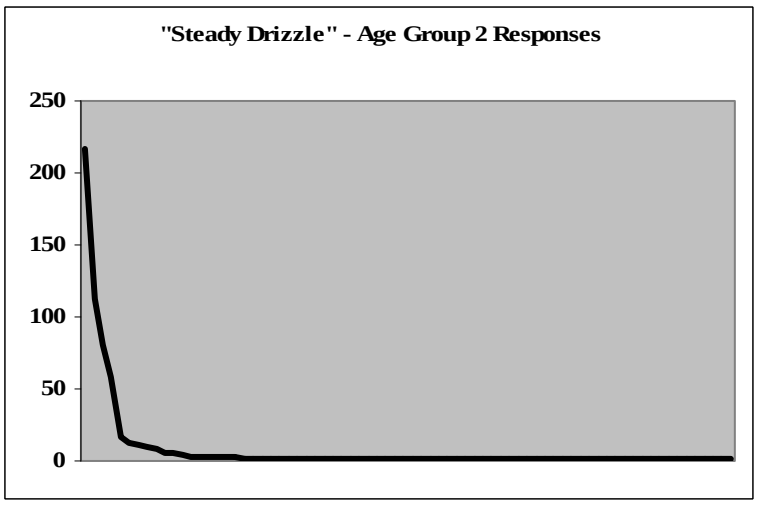
“Steady Drizzle” Frequency Graphs

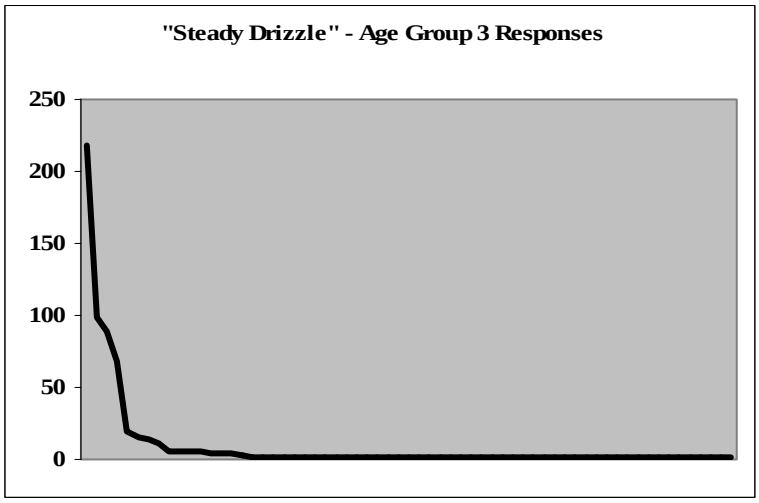


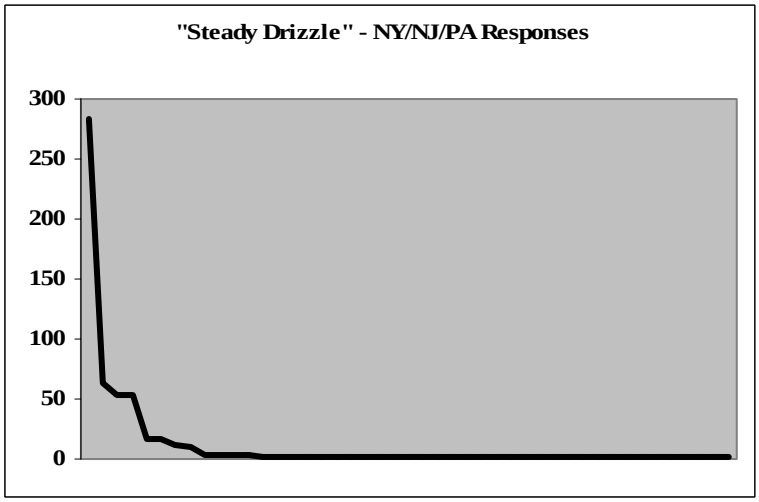


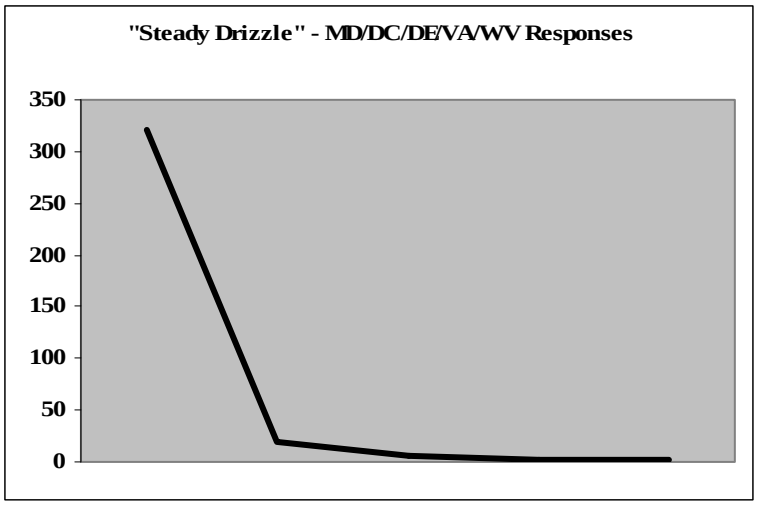

"Steady Drizzle" - Female Responses
300
250
200
150
100
50
0

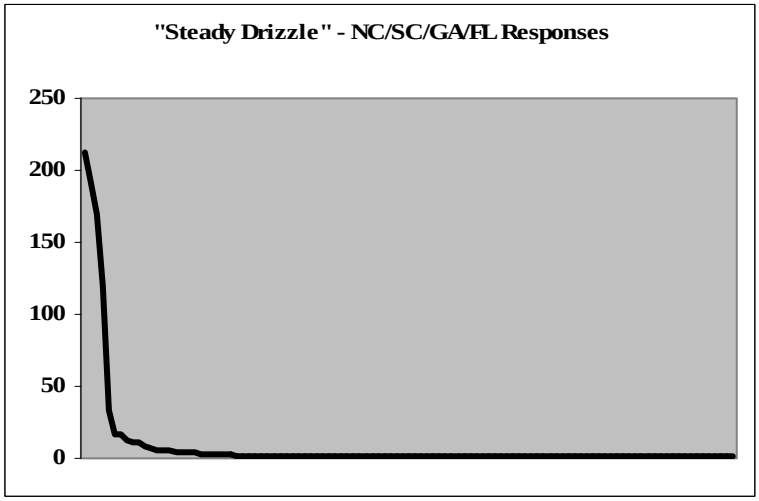

"Steady Drizzle" - Age Group 1 Responses
400
350
300
250
200
150
100
50
0


"Steady Drizzle" - Age Group 2 Responses
250
200
150
100
50
0


"Steady Drizzle" - Age Group 3 Responses
250
200
150
100
50
0


"Steady Drizzle" - NY/NJ/PA Responses
300
250
200
150
100
50
0


"Steady Drizzle" - MD/DC/DE/VA/WV Responses
350
300
250
200
150
100
50
0


"Steady Drizzle" - NC/SC/GA/FL Responses
250
200
150
100
50
0

APPENDIX F
"TOP 10 RESPONSES" CHARTS FOR EACH LEXICAL ITEM

AGRICULTURE/LAND DOMAIN

"Hog Pen" All

hog pen	815
pig pen	272
pen	68
hog lot	64
pig sty	62
hog house	49
hog pasture	40
sty	36
fattening pen	17
floored pen	16

"Hog Pen" Age Group 1

hog pen	297
pig pen	115
pig sty	26
hog lot	16
Pen	10
Sty	10
hog house	9
hog pasture	6
fattening pen	4
hog sty	3

"Hog Pen" Age Group 2

hog pen	241
pig pen	81
hog lot	31
hog house	29
pen	25
pig sty	20
hog pasture	17
sty	11
fattening pen	5
floored pen	4

“Hog Pen” Age Group 3

hog pen	276
pig pen	75
pen	33
hog lot	17
hog pasture	17
pig sty	16
sty	15
hog house	11
floored pen	9
fattening pen	7

“Hog Pen” Males

hog pen	588
pig pen	158
hog lot	48
pen	48
hog house	44
hog pasture	35
pig sty	34
sty	24
fattening pen	14
floored pen	14

“Hog Pen” Females

hog pen	227
pig pen	114
pig sty	28
pen	20
hog lot	16
sty	12
hog house	5
hog pasture	5
hog crawl	4
pin	4

“Hog Pen” New York, New Jersey, Pennsylvania

hog pen	189
pig pen	170
pig sty	41
hog house	29
sty	17
hog yard	4
pen	4
pig house	4
pig yard	4
pig stable	3

“Hog Pen” D.C., Delaware, Maryland, Virginia, West Virginia

- There were only a small number of different responses for this subset, and thus several items with only one or two responses have been included, based on which ones occur first alphabetically.

hog pen	283
pig pen	57
hog house	13
hog lot	11
pig sty	7
hog bed	2
hog pound	2
pen	2
sty	2
hog nest	1

“Hog Pen” North Carolina, South Carolina, Georgia, Florida

hog pen	343
pen	62
hog lot	52
pig pen	45
hog pasture	40
fattening pen	17
sty	17
floored pen	16
pig sty	14
hog crawl	9

AGRICULTURE/LAND DOMAIN

“Meadow” All

meadow	872
swale	128
savanna	56
hayfield	37
meadow land	23
prairie	22
bog	18
pasture	17
savanna land	11
old field	8

“Meadow” Age Group 1

meadow	331
swale	36
savanna	22
hayfield	16
bog	6
pasture	6
bottom	4
prairie	4
lowland	3
canebrake	2

- There were only a small number of different responses for this subset, and thus several items with only two or three responses have been included, based on which ones occur first alphabetically.

“Meadow” Age Group 2

meadow	283
swale	30
savanna	16
meadow land	14
hayfield	10
prairie	8
pasture	6
field	4
bog	3
flat	3

“Meadow” Age Group 3

meadow	255
swale	61
savanna	18
hayfield	11
prairie	10
bog	9
meadow land	8
savanna land	6
swale grass	6
pasture	5

“Meadow” Males

meadow	629
swale	105
savanna	38
hayfield	19
prairie	17
meadow land	15
bog	14
pasture	11
savanna land	9
old field	8

“Meadow” Females

meadow	243
swale	23
hayfield	18
savanna	18
meadow land	8
pasture	6
prairie	5
bog	4
canebrake	3
lowland	3

“Meadow” NY/NJ/PA

meadow	346
swale	128
hayfield	37
swale grass	8
beaver meadow	5
prairie	5
field	4
hay lot	4
meadow land	4
pond	4

“Meadow” MD/DC/DE/VA/WV

meadow	319
meadow land	5
bottom	4
bottom land	2
pasture	2
branch	1
buffalo wallow	1
flats	1
glade	1
grass bottoms	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

"Meadow" NC/SC/GA/FL

meadow	207
savanna	56
bog	18
prairie land	16
pasture	15
meadow land	14
savanna land	11
old field	8
canebrake	7
pasture land	6

AGRICULTURE/LAND DOMAIN

"Swamp" All

Swamp	1186
Marsh	133
Bog	103
swamp land	51
Pond	37
Slough	25
Boggy	24
Bay	23
Branch	18
river swamp	16

"Swamp" Age Group 1

swamp	473
marsh	42
bog	32
branch	8
slough	8
swamp land	8
bay	6
pond	6
slash	5
boggy	4

“Swamp” Age Group 2

swamp	342
marsh	47
bog	34
pond	23
swamp land	14
boggy	13
bay	9
slough	9
mire	8
swampy	8

“Swamp” Age Group 3

swamp	369
marsh	44
bog	37
swamp land	29
bay	8
pond	8
slough	8
swampy land	8
bog hole	7
boggy	7

“Swamp” Males

swamp	824
marsh	86
bog	74
swamp land	39
pond	30
slough	19
bay	17
boggy	16
river swamp	14
fly	13

“Swamp” Females

swamp	362
marsh	47
bog	29
swamp land	12
boggy	8
branch	8
pond	7
bay	6
slough	6
boggy land	5

“Swamp” NY/NJ/PA

swamp	388
bog	30
marsh	23
fly	14
swamp land	11
swampy	7
The Fly	5
slough	4
swampy land	4
bog land	3

“Swamp” MD/DC/DE/VA/WV

swamp	338
marsh	68
bog	6
branch	4
low grounds	2
marshy land	2
mire	2
slash	2
boggy land	1
frog marsh	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Swamp” NC/SC/GA/FL

swamp	460
bog	67
marsh	42
swamp land	40
pond	36
bay	23
boggy	23
slough	20
river swamp	16
hammock	15

FOOD DOMAIN

“Cobbler” All

apple cobbler	130
family pie	115
cobbler	108
apple pie	97
apple dumpling	84
apple pot pie	72
pot pie	65
apple pudding	53
deep dish apple pie	52
peach pie	52

“Cobbler” Age Group 1

apple cobbler	66
family pie	56
cobbler	51
apple pot pie	32
deep dish apple pie	31
apple dumpling	23
pot pie	21
cobbler pie	18
apple pie	16
apple pudding	16

“Cobbler” Age Group 2

apple pie	38
cobbler	37
apple cobbler	32
family pie	27
apple dumpling	21
apple pot pie	21
pie	19
apple pudding	19
apple tart	17
deep dish apple pie	16

“Cobbler” Age Group 3

apple pie	43
apple dumpling	40
apple cobbler	32
family pie	32
pot pie	31
peach pie	30
pie	24
cobbler	20
apple pot pie	19
pan pie	19

“Cobbler” Males

apple cobbler	82
cobbler	74
apple pie	72
apple dumpling	65
family pie	61
apple pot pie	52
pot pie	45
peach pie	43
apple pudding	36
pie	32

“Cobbler” Females

family pie	54
apple cobbler	48
cobbler	34
deep dish apple pie	34
apple pie	25
apple pot pie	20
pot pie	20
apple dumpling	19
deep apple pie	17
apple pudding	17

“Cobbler” NY/NJ/PA

apple dumpling	64
apple pot pie	47
deep dish apple pie	38
apple grunt	37
birds nest	32
apple pudding	31
crows nest	30
pot pie	21
apple cobbler	18
apple pie	15

“Cobbler” MD/DC/DE/VA/WV

apple cobbler	74
cobbler	53
family pie	48
cobbler pie	21
apple pot pie	18
pot pie	16
apple pudding	13
big pie	7
cut and come	5
deep dish apple pie	5

“Cobbler” NC/SC/GA/FL

apple pie	79
family pie	67
peach pie	52
cobbler	49
apple cobbler	38
pie	36
pan pie	34
pot pie	29
potato pie	25
apple tart	24

FOOD DOMAIN

"Cornbread" All

corn bread	743
corn pone	277
johnny cake	198
pone	122
pone of corn bread	56
pone bread	48
pone of bread	47
hoecake	34
bread	32
corn cake	32

"Cornbread" Age Group 1

corn bread	299
corn pone	111
johnny cake	59
pone	27
pone of corn bread	17
corn cake	15
pone of bread	12
hoecakes	9
bread	7
corn muffins	6

"Cornbread" Age Group 2

corn bread	220
corn pone	79
johnny cake	59
pone	51
pone bread	24
pone of corn bread	23
pone of bread	22
bread	16
corn dodger	12
hoecake	12

“Cornbread” Age Group 3

corn bread	223
corn pone	87
johnny cake	79
pone	44
pone bread	20
pone of bread	16
pone of corn bread	16
corn dodger	13
hoecake	13
bread	9

“Cornbread” Males

corn bread	468
corn pone	205
johnny cake	154
pone	100
pone of corn bread	36
pone bread	31
pone of bread	27
corn cake	25
bread	23
corn dodger	19

“Cornbread” Females

corn bread	275
corn pone	72
johnny cake	44
pone of bread	23
pone	22
pone of corn bread	20
pone bread	17
hoecakes	15
corn dodger	11
bread	9

“Cornbread” NY/NJ/PA

johnny cake	189
corn bread	180
corn pone	56
pone	27
corn cake	17
corn meal bread	11
indian bread	10
corn muffins	5
corn meal muffins	4
brown bread	3

- There were only a small number of different responses for this subset, and thus several items with only three or four responses have been included, based on which ones occur first alphabetically.

“Cornbread” MD/DC/DE/VA/WV

corn bread	191
corn pone	136
pone of corn bread	38
pone	30
pone of bread	27
corn dodger	19
pone bread	13
batter bread	10
corn cakes	7
johnny cake	4

“Cornbread” NC/SC/GA/FL

corn bread	372
corn pone	85
pone	65
pone bread	35
bread	29
hoecake	29
pone of bread	20
pone of corn bread	18
corn dodgers	9
corn cake	8

FOOD DOMAIN

“Pancakes” All

pancakes	757
batter cake	358
flitter	304
fritter	260
flapjack	192
flannel cake	180
griddle cake	149
hotcake	76
slapjack	49
wheat cakes	37

“Pancakes” Age Group 1

pancakes	249
batter cake	116
fritter	99
flannel cake	85
flitters	82
flapjack	74
griddle cakes	49
hotcakes	43
wheat cakes	18
cakes	9

“Pancakes” Age Group 2

pancakes	243
batter cake	121
flitters	112
fritters	81
flapjacks	70
flannel cakes	55
griddle cakes	50
hotcakes	23
slapjacks	11
wheat cakes	10

“Pancakes” Age Group 3

pancakes	264
batter cakes	121
flitters	109
fritters	80
griddle cakes	50
flapjacks	48
flannel cakes	40
slapjacks	29
buckwheat cakes	10
hotcakes	10

“Pancakes” Males

pancakes	525
flitters	217
batter cake	210
flapjack	147
fritters	146
flannel cake	122
griddle cakes	92
hotcakes	53
slapjacks	35
wheat cakes	30

“Pancakes” Females

pancakes	232
batter cake	148
fritter	114
flitters	86
flannel cake	58
griddle cakes	57
flapjack	45
hotcakes	23
slapjacks	14
corn fritters	7

“Pancakes” NY/NJ/PA

pancakes	320
griddle cakes	94
flannel cakes	91
flapjacks	64
hotcakes	28
fritters	27
flitters	26
wheat cakes	24
slapjacks	22
buckwheat cakes	11

“Pancakes” MD/DC/DE/VA/WV

pancakes	189
flitters	128
batter cakes	123
fritters	84
flannel cakes	75
flapjacks	56
griddle cakes	22
hotcakes	19
wheat cakes	12
flitter cakes	10

“Pancakes” NC/SC/GA/FL

pancakes	248
batter cake	227
flitters	151
fritters	149
flapjacks	72
griddle cakes	33
hotcakes	29
slapjacks	17
flannel cakes	14
corn fritters	13

HOME/HOUSEHOLD ITEMS DOMAIN

"Andirons" All

andirons	553
firedogs	349
dog irons	258
handirons	116
dogs	73
fire irons	66
grate	18
irons	17
fire rocks	16
fender	5

"Andirons" Age Group 1

andirons	249
dog irons	93
firedogs	91
dogs	21
fire irons	20
handirons	16
grate	4
iron	3
wood dogs	3
fire rocks	2

"Andirons" Age Group 2

andirons	165
firedogs	138
dog irons	70
handirons	33
dogs	31
fire irons	23
fire rocks	8
irons	7
grates	4
dog's irons	2

- There were only a small number of different responses for this subset, and thus several items with only two to four responses have been included, based on which ones occur first alphabetically.

“Andirons” Age Group 3

andirons	135
firedogs	119
dog irons	94
handirons	67
fire irons	23
dogs	21
grates	10
irons	7
fire rocks	6
fenders	3

“Andirons” Males

andirons	333
firedogs	228
dog irons	178
handirons	89
dogs	51
fire irons	46
fire rocks	13
irons	13
grates	11
fenders	5

“Andirons” Females

andirons	220
firedogs	121
dog irons	80
handirons	27
dogs	22
fire irons	20
grates	7
irons	4
fire logs	3
fire rocks	3

“Andirons” NY/NJ/PA

andirons	231
handirons	51
dog irons	21
firedogs	14
dogs	13
irons	10
grate	6
fire irons	5
sadirons	4
fire basket	2

“Andirons” MD/DC/DE/VA/WV

andirons	117
dog irons	102
firedogs	81
handirons	54
fire irons	36
dogs	24
fire rocks	8
wood dogs	4
log irons	3
dog's irons	2

“Andirons” NC/SC/GA/FL

firedogs	254
andirons	205
dog irons	135
dogs	36
fire irons	25
grate	12
handirons	11
fire rocks	7
irons	6
fender	4

HOME/HOUSEHOLD ITEMS DOMAIN

"Hearth" All

hearth	1430
fireplace	202
fire hearth	20
grate	6
hearthrock	5
chimney hearth	4
stove hearth	4
fires	3
hearth stone	3
ash pan	2

- There were only a small number of different responses for this subset, and thus several items with only two or three responses have been included, based on which ones occur first alphabetically.

"Hearth" Age Group 1

hearth	516
fireplace	32
fire hearth	8
grate	3
chimney hearth	2
fires	2
fire screen	1
firerock	1
furnace	1
hat	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

"Hearth" Age Group 2

hearth	452
fireplace	83
fire hearth	6
hearth rocks	5
grate	3
stove hearth	2
ash pan	1
brick hearth	1
clay	1
clay hearth	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Hearth” Age Group 3

hearth	456
fireplace	86
fire hearth	6
chimney hearth	2
hearth stone	2
stove hearth	2
ash pan	1
ash pit	1
clay	1
coal grates	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Hearth” Males

hearth	953
fireplace	138
fire hearth	16
hearth rocks	5
chimney hearth	4
grate	4
stove hearth	3
clay	2
fire	2
hearth stone	2

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Hearth” Females

hearth	477
fireplace	64
fire hearth	4
grate	2
ash pan	1
cricket on the hearth	1
fender in front	1
fenders	1
fireplace hearth	1
fires	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Hearth” NY/NJ/PA

hearth	398
fireplace	56
fire hearth	4
stove hearth	4
ash pan	2
ash pit	1
dog	1
elevated oven	1
false fireplace	1
fire	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Hearth” MD/DC/DE/VA/WV

Hearth	392
fire hearth	6
Hearthrock	2
Firerock	1
hearth stone	1
kindling Holz	1

- For this subset, there were not ten unique responses given; thus, only the six different responses were included, resulting in the inclusion of a few items with only one response.

“Hearth” NC/SC/GA/FL

Hearth	640
Fireplace	146
fire hearth	10
Grate	6
chimney hearth	4
hearth rocks	3
Clay	2
Fires	2
rock hearth	2
brick hearth	1

HOME/HOUSEHOLD ITEMS DOMAIN

"Pallet" All

pallet	652
bunk	165
lodging	38
mattress	30
featherbed	24
trundle bed	19
lodge	12
bed on the floor	10
shakedown	6
bed	4

"Pallet" Age Group 1

pallet	218
Bunk	80
Lodging	16
Mattress	4
bed on the floor	3
Pad	2
Shakedown	2
Spread	2
Bed	1
Cot	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

"Pallet" Age Group 2

Pallet	215
Bunk	51
Lodging	16
trundle bed	8
Featherbed	7
Mattress	6
Lodge	4
Bed	3
bed on the floor	3
Shakedown	3

“Pallet” Age Group 3

Pallet	218
bunk	34
mattress	20
featherbed	17
trundle bed	10
lodge	7
lodging	6
bed on the floor	4
straw tick	2
bedsteads	1

- There were only a small number of different responses for this subset, and thus several items with only one or two responses have been included, based on which ones occur first alphabetically.

“Pallet” Males

pallet	391
bunk	148
mattress	24
lodging	23
featherbed	19
trundle bed	11
lodge	8
bed on the floor	7
shakedown	4
spread	3

“Pallet” Females

pallet	261
bunk	17
lodging	15
trundle bed	8
mattress	6
featherbed	5
lodge	4
bed on the floor	3
pad	3
bed	2

- There were only a small number of different responses for this subset, and thus several items with only two or three responses have been included, based on which ones occur first alphabetically.

“Pallet” NY/NJ/PA

bunk	104
featherbed	9
shakedown	5
straw tick	3
bollick	2
cord bedstead	2
feather tick	2
mattress	2
bed	1
floor bed	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Pallet” MD/DC/DE/VA/WV

Pallet	244
Bunk	56
Lodge	12
Spread	4
bed on the floor	2
floor bed	1
Pad	1

- For this subset, there were not ten unique responses given; thus, only the seven different responses were included, resulting in the inclusion of a few items with only one response.

“Pallet” NC/SC/GA/FL

Pallet	407
Lodging	38
mattress	28
trundle bed	18
featherbed	15
bed on the floor	8
bunk	5
bed	3
made-down bed	2
on the floor	2

- There were only a small number of different responses for this subset, and thus several items with only two responses have been included, based on which ones occur first alphabetically.

WEATHER DOMAIN

“Cloudburst” All

Cloudburst	345
Downpour	332
heavy rain	171
hard rain	146
gully washer	119
big rain	84
Flood	84
Pourdown	61
hard shower	57
heavy shower	47

“Cloudburst” Age Group 1

Downpour	172
Cloudburst	163
heavy rain	54
hard rain	41
gully washer	33
hard shower	20
big rain	18
pourdown	16
flood	15
heavy shower	12

“Cloudburst” Age Group 2

downpour	97
cloudburst	90
heavy rain	51
hard rain	49
gully washer	43
big rain	33
flood	31
pourdown	24
hard shower	22
heavy shower	17

“Cloudburst” Age Group 3

cloudburst	91
heavy rain	66
downpour	63
hard rain	56
gully washer	42
flood	38
big rain	33
pourdown	21
heavy shower	18
hard shower	15

“Cloudburst” Males

cloudburst	260
downpour	181
heavy rain	132
hard rain	103
gully washer	82
flood	68
big rain	56
pourdown	43
heavy shower	39
hard shower	33

“Cloudburst” Females

downpour	151
cloudburst	85
hard rain	43
heavy rain	39
gully washer	37
big rain	28
hard shower	24
pourdown	18
flood	16
shower	9

“Cloudburst” NY/NJ/PA

cloudburst	140
downpour	94
heavy rain	41
heavy shower	32
hard rain	23
flood	21
hard shower	21
pouring rain	12
pourdown	10
big rain	7

“Cloudburst” MD/DC/DE/VA/WV

downpour	94
cloudburst	77
hard rain	47
heavy rain	43
gully washer	29
pourdown	27
hard shower	22
big rain	17
flood	13
pouring down rain	8

“Cloudburst” NC/SC/GA/FL

downpour	144
cloudburst	128
gully washer	88
heavy rain	87
hard rain	76
big rain	60
flood	50
trash mover	35
pourdown	24
squall	17

WEATHER DOMAIN

"Dry spell" All

Dry spell	906
drouth	720
drought	164
dry weather	38
dry time	25
Dry drouth	15
dry drought	14
dry season	10
wet spell	9
spell of dry weather	8

"Dry spell" Age Group 1

dry spell	324
drouth	232
drought	67
dry weather	9
dry season	3
spell of dry weather	3
dry drought	2
dry time	2
dry	1
dry drouth	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

"Dry spell" Age Group 2

dry spell	281
Drouth	243
Drought	64
dry weather	15
dry drought	7
dry time	6
dry drouth	5
wet spell	5
Dry	3
dry season	3

- There were only a small number of different responses for this subset, and thus several items with only three responses have been included, based on which ones occur first alphabetically.

“Dry spell” Age Group 3

dry spell	298
drouth	242
drought	33
dry time	17
dry weather	14
dry drouth	9
dry drought	5
long dry spell	4
dry season	4
wet spell	3

“Dry Spell” Males

dry spell	640
drouth	547
drought	88
dry weather	25
dry time	12
dry drought	10
dry drouth	10
wet spell	8
long dry spell	6
dry season	5

“Dry spell” Females

dry spell	266
drouth	173
drought	76
dry time	13
dry weather	13
dry drouth	5
dry season	5
dry drought	4
spell of dry weather	3
dry	2

“Dry spell” NY/NJ/PA

dry spell	310
drouth	249
drought	58
dry time	11
spell of dry weather	7
dry weather	6
dry season	3
long dry spell	1
a spell of weather	1
clear spell	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Dry spell” MD/DC/DE/VA/WV

dry spell	265
drouth	262
drought	28
dry weather	10
dry time	7
dry drouth	2
dry season	2
draft	1
dry	1
dry drought	1

- There were only a small number of different responses for this subset, and thus several items with only one response have been included, based on which ones occur first alphabetically.

“Dry spell” NC/SC/GA/FL

dry spell	331
drouth	209
drought	78
dry weather	22
dry drought	13
dry drouth	13
wet spell	8
dry time	7
long dry spell	6
dry season	5

WEATHER DOMAIN

“Steady drizzle” All

steady drizzle	796
drizzle	286
shower	233
sprinkle	173
mist	51
light rain	27
steady rain	27
light shower	26
steady	17
misting rain	12

“Steady drizzle” Age Group 1

steady drizzle	361
drizzle	74
shower	62
sprinkle	46
mist	15
light rain	4
drizzling rain	3
drizzly	3
misting rain	3
steady	3

“Steady drizzle” Age Group 2

steady drizzle	217
drizzle	113
shower	81
sprinkle	58
mist	17
light rain	12
steady rain	11
light shower	10
steady	9
drizzling rain	5

“Steady drizzle” Age Group 3

steady drizzle	218
drizzle	98
shower	89
sprinkle	68
mist	19
light shower	15
steady rain	14
light rain	11
spitting	6
little shower	5

“Steady drizzle” Males

steady drizzle	542
drizzle	198
shower	165
sprinkle	122
mist	36
light shower	19
steady rain	19
light rain	18
steady	16
misting rain	8

“Steady drizzle” Females

steady drizzle	254
drizzle	88
shower	68
sprinkle	51
mist	15
light rain	9
steady rain	8
light shower	7
drizzling rain	6
drizzly	4

“Steady drizzle” NY/NJ/PA

steady drizzle	284
shower	64
drizzle	54
sprinkle	53
mist	17
steady rain	16
light rain	11
light shower	10
little shower	4
steady	4

“Steady drizzle” MD/DC/DE/VA/WV

steady drizzle	321
drizzling	20
spitting	6
splitting snow	1
steady grizzle	1

- *drizzling* is the only variation of *drizzle* to appear in this set
- For this subset, there were not ten unique responses given; thus, only the five different responses were included, resulting in the inclusion of a few items with only one response.

“Steady drizzle” NC/SC/GA/FL

drizzle	212
steady drizzle	191
shower	169
sprinkle	120
mist	34
light rain	16
light shower	16
steady	13
misting rain	11
steady rain	11

WORKS CITED

Brown, Nina. "Hans Kurath: Linguistic Atlas of the United States." CSISS Classics. 2009. Center for Spatially Integrated Social Science. 13 Mar 2009 <http://www.csiss.org/classics/content/17>.

Burkette, Allison. "The Story of Chester Drawers." American Speech 76(2001): 139-157.

Butters, Ronald R. "Chance as Cause of Language Variation and Change." Journal of English Linguistics 29(2001): 201-213.

Bybee, Joan. Frequency of Use and the Organization of Language. New York: Oxford University Press, 2007.

Bybee, Joan and Paul Hopper. "Introduction to Frequency and the Emergence of Linguistic Structure." Frequency and the Emergence of Linguistic Structure. Ed. Joan Bybee and Paul Hopper. Philadelphia: John Benjamins Publishing Co., 2001.

Cameron, Lynne, and Diane Larsen-Freeman. "Complex Systems and Applied Linguistics." International Journal of Applied Linguistics 17(2007): 226-240.

Density Estimate Map: andirons. Map. University of Georgia. 25 Mar 2009 <http://us.english.uga.edu/cgi-bin/lapsite.fcgi/lamsas/de-maps/andirons_8_3.html>.

Density Estimate Map: dog irons. Map. University of Georgia. 25 Mar 2009 <http://us.english.uga.edu/cgi-bin/lapsite.fcgi/lamsas/de-maps/dog_irons_8_3.html>.

Density Estimate Map: firedogs. Map. University of Georgia. 25 Mar 2009 <http://us.english.uga.edu/cgi-bin/lapsite.fcgi/lamsas/de-maps/firedogs_8_3.html>.

Density Estimate Map: handirons. Map. University of Georgia. 25 Mar 2009 <http://us.english.uga.edu/cgi-bin/lapsite.fcgi/lamsas/de-maps/handirons_8_3.html>.

"Drought, Drouth." Oxford English Dictionary. 1989. Oxford University Press. 15 Mar 2009 <http://dictionary.oed.com/cgi/entry/50070155?single=1&query_type=word&queryword=drouth&first=1&max_to_show=10>.

Hoover, Sandra. "Lexical Variation and Change in Farming Words: 1970-2001." MA Thesis. University of Georgia, 2001. 04 Apr. 2009 < http://dbs.galib.uga.edu/cgi-bin/write_stats.cgi?stattype=fulltext&dbscode=getd&format=pdf&redirect=http://getd.galib.uga.edu/campus/hoover_sandra_e_200112_ma/hoover_sandra_e_200112_ma.pdf>

Johnson, Ellen. Lexical Change and Variation in the Southeastern United States, 1930-1990. Tuscaloosa: University of Alabama Press, 1996.

Kretzschmar, William A., Jr, and Susan Tamasi. "Distributional Foundations for a Theory of Language Change." World Englishes 22(2003): 377-401.

Kretzschmar, William A., Jr. Dialectology and Complex Systems. Methods in Dialectology Conference. Leeds: 2008.

Kretzschmar, William A., Jr. "Neural networks and the linguistics of speech." Interdisciplinary Science Reviews 33(2008): 336-356.

Kretzschmar, William A., Jr. The Linguistics of Speech. Cambridge: Cambridge University Press, 2009.

Kurath, Hans. A Word Geography of the Eastern United States. Ann Arbor: University of Michigan, 1949.

"LAMSAS." Linguistic Atlas Projects. 26 Mar 2005. University of Georgia. 12 Mar 2009 <http://us.english.uga.edu/cgi-bin/lapsite.fcgi/lamsas/>.

Larsen-Freeman, Diane, and Lynne Cameron. "Research Methodology on Language Development from a Complex Systems Perspective." Modern Language Journal 92(2008): 200-213.

"Linguistic Atlas Projects." Linguistic Atlas Projects. 10 June 2005. University of Georgia. 22 Feb 2009 <http://us.english.uga.edu/>.

Sapir, Edward. Language: An Introduction to the Study of Speech. New York: Harcourt, 1921.