Digital Archive of Southern Speech (DASS): Introduction

William A. Kretzschmar, Jr.
University of Georgia

The Linguistic Atlas Project

The Linguistic Atlas Project (LAP) consists of a set of survey research projects about the words and pronunciation of everyday American English, the largest project of its kind in the country. Beginning in 1929, interviews with thousands of native speakers have been carried out in different regions of the country--and primary interviews are still going on in the Western States. It is not hard to see why the work has taken so long when each survey subject has literally spent hours in conversation with a field worker, talking about common topics like family, the weather, household articles and activities, agriculture, and social connections. A total of over 800 topics were covered in each interview in most regions, to elicit the common words in use for them. Before tape recorders, each response was recorded by highly-trained field workers in detailed phonetic script to preserve the pronunciation. Audio tapes later allowed phonetic transcription in the office rather than in the field. The regional surveys have revealed a tremendous wealth of language variation in American English, both in vocabulary and pronunciation. A large part of the continental US has been surveyed, all of the US east of the Mississippi, but most of the results of the interviews have not yet been published in an accessible way. All surviving LAP audio recordings (nearly 7000 hours, from the Gulf States, the North-Central States, and the Western States) will be preserved indefinitely in digital computer storage, with metadata and finding aids for the recordings on the existing LAP Web site.

Recordings to be preserved and offered to the public come from different regions. The Linguistic Atlas of the Gulf States (LAGS) surveyed the everyday speech of Georgia, Tennessee, Florida, Alabama, Mississippi, Arkansas, Louisiana, and Texas in a series of 914 audio-taped interviews in the late 1960s and 1970s. Interviews average about 6 hours in length. The systematic LAGS tape archive amounts to 5500 hours of sound recordings. The major collections besides LAGS come from the North-Central States, including Roger Shuy's interviews in Northern Illinois and Lee Pederson's interviews in metropolitan Chicago from the 1960s. These interviews were conducted with methods compatible with LAGS, and thus constitute invaluable comparative evidence from another region of the country. About 50 interviews were conducted in the Western States (Colorado, Utah, Wyoming) in the 1980s; 50 more interviews have been conducted more recently in California, Colorado, and El Paso, TX. The Western States interviews use a new three-hour interview plan compatible with that of the older interviews. Future LAP interviews can be added to the framework created, so that all may be preserved for posterity in a common resource.

The importance of the LAP recordings is both linguistic and cultural. There is no better systematic coverage of twentieth-century daily life in America than these recordings. The content of the interviews talks about American history and culture according to a predictable set of topics, so that speakers and regions can be compared to one another. The cultural content comes not as dry lists of facts but instead in the original voice of the speakers, culture as it was lived by real people. While previous publications of Atlas data have been valuable for the study of language variation, these interview recordings have great humanistic value as a record of the daily lives of Americans. And the recordings do still have value for linguists, more value than ever since the sound files can be processed with the contemporary computer methods of acoustical phonetics (Erik Thomas, for instance, as made great use of such historical recordings for acoustic analysis, e.g. in Bailey and Thomas 1998). Moreover, since it does not require special training to listen to recordings (as it does to make good use of phonetic transcriptions), members of the public will have much greater access to real American voices, as they were collected systematically by LAP.

It is fair to say that LAP research underlies a good deal of what we know about the twentieth-century varieties of American English. The LAP surveys are the empirical benchmark against which linguists of many kinds can measure their own current research. Earlier analytical volumes based on partial LAP survey results (Kurath 1949, Atwood 1953, Kurath and McDavid 1961) still form the basis for scholarly opinion about

American dialects. These analyses have been confirmed by later work, e.g. Carver 1987, which was based on the research for the *Dictionary of American Regional English* (DARE; Cassidy and Hall 1985-). DARE, of course, is intended to collect all words with regional or “folk” distribution, and to present them in dictionary format. LAP on the other hand, while its results have made a strong contribution to DARE, has a much finer network of survey participants, and is designed to show regional and social variation among them in much greater detail for a more limited, everyday set of topics. LAP survey research has also been a significant starting point for the work of William Labov, who has cited LAP results in many of his major works on American varieties (1963, 1966, 1994); his results as published in such works as his 1991 article on “The Three Dialects of English” (and also his formerly unpublished Telsur results, now available in the Labov, Boberg, and Ash 2006 "Phonological Atlas") confirm in large part and extend earlier LAP analyses of the Midland region (generally corresponding to his area of low-back vowel merger), the Northern region (generally corresponding to the area of the Northern Cities Shift), and Southern region (generally corresponding to the area of the Southern Shift). The Telsur/Phonological Atlas study selected only a few people from each metropolitan district in the country (usually two), in contrast to the dense regional survey pattern of LAP, and so it cannot make possible the detailed regional analyses that have characterized LAP. Labov’s other research in particular cities like Philadelphia, on the other hand, is more dense and community-oriented than the LAP surveys--but the data from these studies, like his earlier studies, remains unavailable to other researchers or the public, who glimpse it only through his analyses. It is fair to say that the historical surveys of LAP can provide a standard against which current research like Labov’s can measure phonological change in large terms (as in the Phonological Atlas) or can assess local community language variation in a wider context; it is also fair to say that, without the results of the historical LAP surveys, scholars would find it much more difficult to assess the value of current research based either on dense local interviews or on a loose network of national interviews.

The existing LAP Web site (http: //www.lap.uga.edu) has become the most accessible source of information about regional American English for the general public as well as researchers. The scholarly works mentioned above, including earlier LAP analyses, have been largely technical in nature, best suited for specialists. DARE is much more accessible to the general public in its content, and it has sold thousands of copies to libraries, but unlike the LAP Web site it is not available in the millions of American homes that now regularly access the Internet. The LAP Web site has been used for teaching American English both in North America and worldwide, and of those who access the site a very large number come from non-academic Internet addresses. Many of the speech patterns documented by LAP still exist, so public and expert users can look up words and pronunciations that they have noticed in their own speech in the highly-interactive portions of the site for Middle and South Atlantic (LAMSAS) and for African-American and Gullah speakers. It is certainly true, however, that interviews carried out in the 1930s and 1940s, even interviews from the 1970s, are better used for historical purposes than for descriptions of contemporary speech. The LAP surveys systematically document the roots of American English, not only from earlier in this century but extending back into the nineteenth century as well. Many of the speakers who were interviewed in the 1930s were old, and had first learned their native tongue as far back as the Civil War. The interviews, however, were conducted with speakers of different ages and social circumstances, in cities as well as rural areas, including interviews with African Americans, so the LAP surveys include a wide spectrum of American speech. Thus, a great many Americans can use the LAP Web site to find out about American words that are still relevant to them and reflect their American cultural heritage.

Language variation is of perennial interest to the general public, and also to different professional fields. The author is frequently called for comment by the press, from William Safire when he used to write his “On Language” column in the *New York Times* (more recently for a *Times* article on the speech of Mayor Bloomberg) to reporters for local papers and radio stations. For historians of culture, LAP data is a treasure trove: Johnson 1996, which compares LAP interviews from the 1930s with her own interviews from 1990, is able precisely to document cultural change over a 50-year span as shown in Southeastern vocabulary. For instance, Johnson found that there is more variation in Southeastern vocabulary now than there was in the 1930s; while the culture has changed and people no longer recognize words for outmoded agricultural implements and horse-drawn vehicles, people have at the same time developed a richer vocabulary for other aspects of daily life. Linguists of many kinds benefit from LAP, such as lexicographers, sociolinguists, and historians of the language. The author and his assistants have produced a series of works (e.g. Kretzschmar 1992, 1996a,b,c; Kretzschmar and Light 1996; Kretzschmar and Konopka 1996; Kretzschmar and Schneider 1996) which bring LAP empirical evidence to bear upon basic questions about the areal and social distribution of words, which in turn raise questions of interest for linguistic theory (e.g. Kretzschmar and Tamasi 2003, Kretzschmar 2009). Most recently, Atlas data has been the foundation for advanced statistical studies such as neural network analysis (Kretzschmar 2004) and Levenstein distance measurements with MDS (Nerbonne 2004). Besides linguists, historians use our data, like the LSU professor who requested that we prepare for him our data on names for the Civil War. Educators are interested because it helps them to know about students' home language, thus differences from the school English they are trying to teach. Speech pathologists need our data so that they can help people with speech problems sound like their neighbors. The National Fair Housing Association is interested in discrimination on the basis of how people sound when they phone realtors ("linguistic profiling"), and our data can assist in that work (see Kretzschmar In Press b).

The LAP audio archive is an unparalleled resource for study not only of the common language of our country, but for American culture more generally. Study of LAP interviews so far has taken advantage only of small bits of transcribed data extracted from the full interview. The availability of complete interviews will make possible the full range of type/token analysis normally carried out by sociolinguists, including more detailed assessment of syntax not possible from the written records. The phonetic transcriptions, as valuable as they have been, have not told the stories of daily life in America in which the extracted terms were embedded. LAP interviews did resort to question-and-answer styles, but the ideal for LAP field work has always been “shotgun” questions, by which conversation was directed towards targeted items, and so the large, untranscribed bulk of the interviews normally consists of the speakers’ accounts of their lives and their families, their occupations and their diversions, their houses and their land. The LAP team has always appreciated the personal, individual nature of each interview, as well as the way that the interviews can represent American culture.

Digital Archive of Southern Speech

The *Linguistic Atlas of the Gulf States* (commonly known as LAGS; Pederson 1986-92) is one of the monumental achievements of twentieth-century linguistics, and yet access to its hundreds of interviews with speakers from Florida to Tennessee to Texas has so far been limited to written transcriptions (whether in print or digital form). The Digital Archive of Southern Speech (DASS), now available on external USB drives from the Linguistic Atlas Project Office, is a collection of 64 full interviews selected by Lee Pederson to cover the range of speech in the Gulf States.

Figure 1 "Grid Units of the LAGS Territory," LAGS vol. 1, Pederson 1986-92.

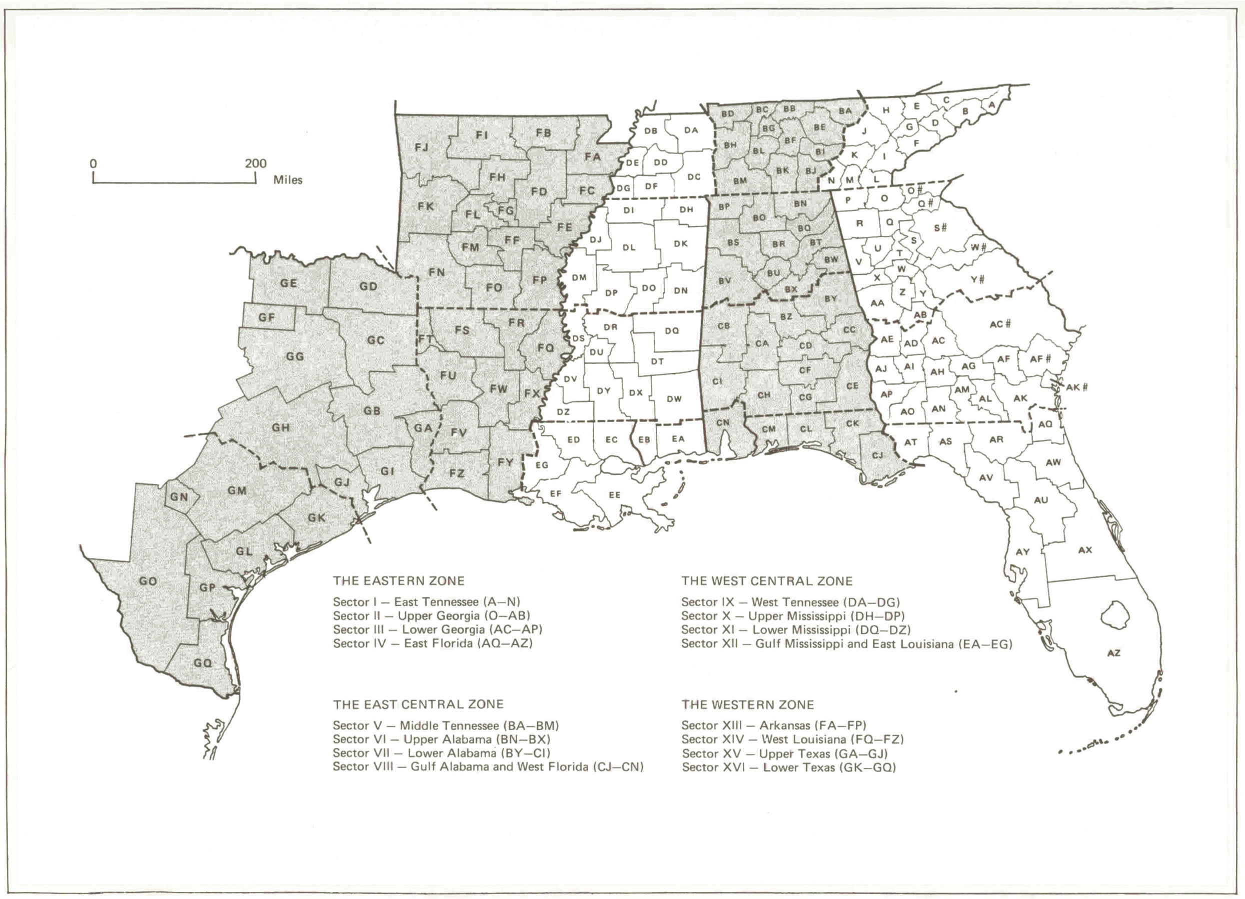
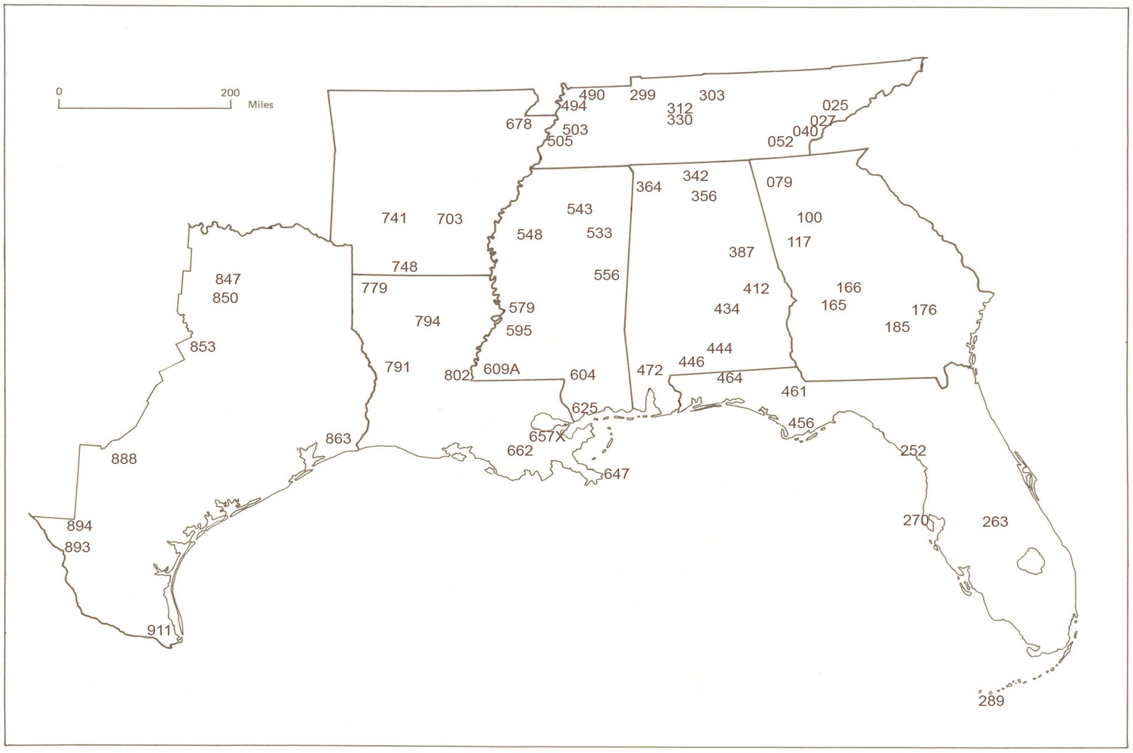

0
200
Miles
THE EASTERN ZONE
Sector I – East Tennessee (A–N)
Sector II – Upper Georgia (O–AB)
Sector III – Lower Georgia (AC–AP)
Sector IV – East Florida (AQ–AZ)
THE WEST CENTRAL ZONE
Sector IX – West Tennessee (DA–DG)
Sector X – Upper Mississippi (DH–DP)
Sector XI – Lower Mississippi (DQ–DZ)
Sector XII – Gulf Mississippi and East Louisiana (EA–EG)
THE EAST CENTRAL ZONE
Sector V – Middle Tennessee (BA–BM)
Sector VI – Upper Alabama (BN–BX)
Sector VII – Lower Alabama (BY–CI)
Sector VIII – Gulf Alabama and West Florida (CJ–CN)
THE WESTERN ZONE
Sector XIII – Arkansas (FA–FP)
Sector XIV – West Louisiana (FQ–FZ)
Sector XV – Upper Texas (GA–GJ)
Sector XVI – Lower Texas (GK–GQ)

Figure 1 Grid Units of the LAGS Territory

First, what interviews are included? DASS constitutes a quota sample of LAGS speakers. First, there are four speakers from each of the sixteen regional divisions of LAGS (Figure 1), the four east-west zones each further subdivided four ways north-south into sectors. The actual locations of the interviews cover the area but are not spread exactly evenly, as Figure 2 shows. This is partly because Lee Pederson chose the interviews that he thought were best from each sector, and also because we were also applying social quotas, not merely geography.

Figure 2 Locations of DASS speakers

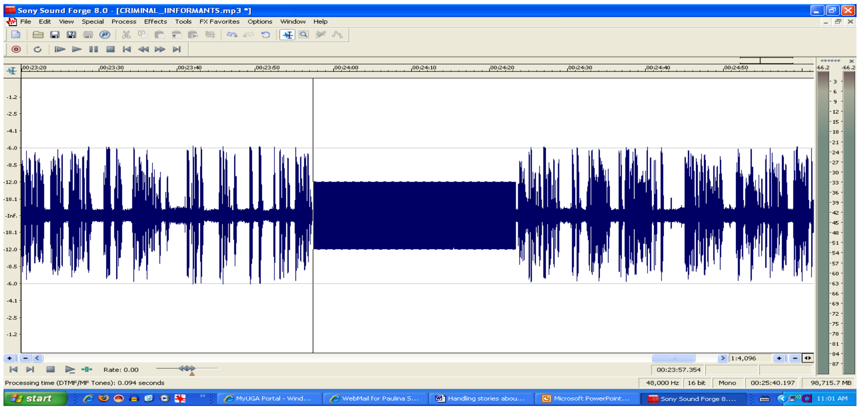


Of the four people per sector, Pederson selected one non-African American from each of the Atlas speaker types: type one are folk speakers, who are older and less educated; type two are common speakers, usually with high-school education; and type three are cultivated speakers, usually with at least some college or compensating experience. The fourth speaker per sector was African American, and the speaker types of African Americans are balanced across the sixteen sectors. Speakers range in age from 15 to 90, with an average age of 61. There are 30 women and 34 men. The characteristics of each speaker are listed in a spreadsheet that comes with DASS. For those interested in

Pederson's innovative correlation of land areas in the LAGS region with language, we provide land region as one of the speaker characteristics in the spreadsheet, and also provide Pederson's map of land regions. So, the DASS sample of LAGS covers the geographic and social range of the speakers interviewed.

DASS includes digital audio files in two kinds, WAV format (long, uncompressed files, useful for acoustic phonetic or other technical processing) and MP3 (small, compressed files good for listening). The interviews were originally recorded in the field on reel-to-reel audio tape. We made a digital version of every reel of tape, one WAV file per reel, usually about an hour of sound. Each interview thus consists of a set of from 3 to 13 reels, from 3 to 13 hours of interview. We have kept an unaltered archival copy of every WAV file, but we make available in DASS a "public" version of the file. Personally identifying or sensitive information in the files was replaced with a tone, or as we say "beeped out," to protect the privacy and to assure ethical treatment of speakers. Figure 3 shows the waveform display of an interview at the point where a "beep" was inserted: the heavy bar that replaces the spiky waves of normal speech in one section.

Figure 3 Waveform display of "beeping"

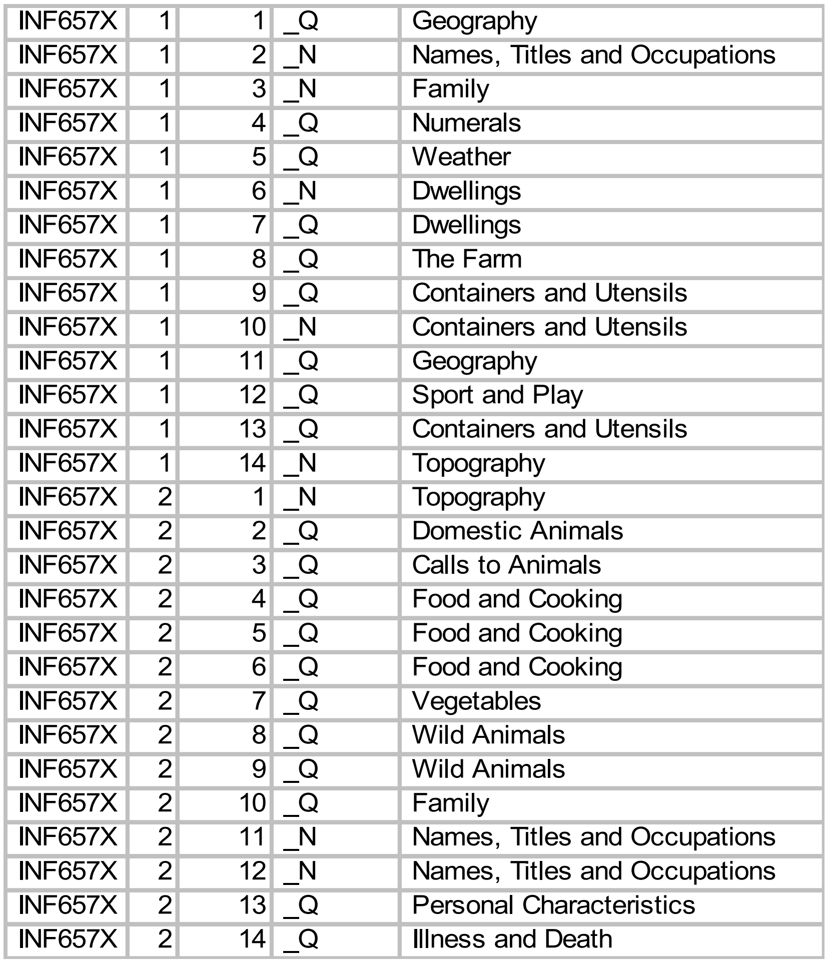


Sony Sound Forge 8.0 - [CRIMINAL_IINFORMANTS.mp3 *]

Processing time (DTMF/MF Tones): 0.094 seconds

No human subjects guidelines like those in place today were in use in linguistics at the time of the interviews in the 1970s and 1980s; in the terms of Institutional Review Board policy, these archival recordings are exempt from current guidelines. Even so, we want to make them available within the spirit of modern practices for the protection of research subjects, and thus the "beeps". Aside from such privacy protection, we have tried to offer as detailed and as true a sound signal as possible. In a few cases we have had to try to enhance the quality of the sound to make it listenable. However, we have generally avoided applying filters to the sound, which was digitized at the 48 Khz sampling rate and 24 bit depth in common use for linguistic interviews. We do think that in most cases it will be possible to extract acoustic phonetic formants from the sound files (as, for example, Erik Thomas has done with LAGS recordings previously, cited above).

Figure 4 Partial list of MP3 files with topics from Speaker 657X



INF657X	1	1	_Q	Geography
INF657X	1	2	_N	Names, Titles and Occupations
INF657X	1	3	_N	Family
INF657X	1	4	_Q	Numerals
INF657X	1	5	_Q	Weather
INF657X	1	6	_N	Dwellings
INF657X	1	7	_Q	Dwellings
INF657X	1	8	_Q	The Farm
INF657X	1	9	_Q	Containers and Utensils
INF657X	1	10	_N	Containers and Utensils
INF657X	1	11	_Q	Geography
INF657X	1	12	_Q	Sport and Play
INF657X	1	13	_Q	Containers and Utensils
INF657X	1	14	_N	Topography
INF657X	2	1	_N	Topography
INF657X	2	2	_Q	Domestic Animals
INF657X	2	3	_Q	Calls to Animals
INF657X	2	4	_Q	Food and Cooking
INF657X	2	5	_Q	Food and Cooking
INF657X	2	6	_Q	Food and Cooking
INF657X	2	7	_Q	Vegetables
INF657X	2	8	_Q	Wild Animals
INF657X	2	9	_Q	Wild Animals
INF657X	2	10	_Q	Family
INF657X	2	11	_N	Names, Titles and Occupations
INF657X	2	12	_N	Names, Titles and Occupations
INF657X	2	13	_Q	Personal Characteristics
INF657X	2	14	_Q	Illness and Death

Besides WAV files, we also offer MP3 versions of the interviews. While each WAV file is very large and long, each MP3 file is small and lasts only about 4.5 minutes. While they are listening for sensitive information to beep out, our assistants record the topic of conversation about at 4.5 minute intervals; these short segments are then automatically exported as separate MP3 file, named for the topic they start with. Figure 4 shows a portion of the list of the MP3 files and topics from the first two reels of a New Orleans interview. Each full reel has 14 MP3 files made from it. Topics come from our set list of 40 designations, drawn from the arrangement of the questionnaire used to guide the interviews. We do not always have long questions and short answers, however. As the list shows, some MP3 files, quite a few actually, are marked with an "N" for narrative, which indicates that there is at least one minute of continuous talk by the speaker in the file. Yes, the interviews have stretches of formal elicitation, but even segments not specially marked as containing narrative can be quite conversational. Overall, DASS contains no fewer than 4943 different MP3 files.

Figure 5 Sample metadata file (speaker 657X)

I. Informant number: 657X
II. Reel number: 1
III. Student's Name: Andrew Paczkowski
IV. Date started: 09.15.2008
V. Date finished: 09.16.2008
VI. Actions taken in dealing with the reel.

1. Reel features: 48,000 Hz, 24 bit, Mono.
2. Tempo edited to 66.500 beats per minute
3. Auto region set up through "Built regions using the current tempo." Tempo set by 60 measures and 60 beats.
4. Reel saved as LAGS(INF657X)1.wav in the C:\Andy\LAGS(INF657X)\LAGS(INF657X)1 folder
5. Beep saved as 'Dial string: 2; 'Single tone length': depending on selection (seconds); 'Break length': 0.001; 'Pause length': 5.000
6. 00:00:24.704-00:00:30.122 deleted. Beep inserted. Length of beep 00:00:05.418
7. 00:00:32.341-00:00:35.075 deleted. Beep inserted. Length of beep 00:00:02.734
8. 00:05:51.962-00:05:52.633 deleted. Beep inserted. Length of beep 00:00:00.670
9. 00:05:57.088-00:05:59.328 deleted. Beep inserted. Length of beep 00:00:02.240
10. 00:42:38.864-00:42:39.226 deleted. Beep inserted. Length of beep 00:00:00.362
11. 00:51:57.568-00:51:58.005 deleted. Beep inserted. Length of beep 00:00:00.436
12. Regions extracted from LAGS(INF657X)1.wav and saved as wav files in C:\Andy\LAGS(INF657X)\LAGS(INF657X)1\LAGS(INF657X)1wav
13. Files from LAGS(INF657X)1wav converted into mp3 format and saved in C:\Andy\LAGS(INF657X)\LAGS(INF657X)1\LAGS(INF657X)1mp3
14. Folder with wav files deleted.
15. Informant: comfortable
16. Interview: formal
17. Additional notes: auxiliary interviewer speaks from ~00:42:12-00:42:28

The last kind of information we offer in DASS is metadata, that is, information about the speakers, about their interviews, and about how their files were processed. We include as a text file the biography for each of the DASS speakers that was originally published in LAGS, with permission of University of Georgia Press. We also include text files that describe exactly what was done to each reel. In Figure 5 we see at the bottom the judgment of the assistant about the quality of the interview, comfortable and formal, and also special features of the recording like the presence of talk from an auxiliary speaker. Our digital files are made from old tapes, long past their usable lifetimes, and yet we have few cases where we cannot extract listenable sound. Finally, we provide spreadsheets that provide information about the speakers, about the labels we have used for topics, and about the sound files provided.

Some users will be happy just to have the two kinds of sound files, and the text, map, and spreadsheet information about the speakers. However, we are pleased to offer a user interface with the files that should make DASS much more usable for a wide range of users. A customized version of LICHEN software comes with DASS on a separate CD, with its own introductory materials.

LAP Staff for DASS

William A. Kretzschmar, Jr., Editor in Chief.
Debbie Vaughn, Research Manager
Paulina Bounds, Lead Graduate Assistant
Steven Coats, Graduate Assistant
Tony Snodgrass, Graduate Assistant
A large team of part-time undergraduate assistants.

Bibliography

(including references to historical and analytical works in language variation)

Allen, Harold. 1973-76. *Linguistic Atlas of the Upper Midwest*. Minneapolis: U. of Minnesota Press.

Atwood, E. Bagby. 1953. *A Survey of Verb Forms in the Eastern United States*. Ann Arbor: U. of Michigan Press.

Bailey, Guy, and Erik Thomas. 1998. Some Aspects of AAVE Phonology. In Mufwene, Salikoko, John Rickford, Guy Bailey, and John Baugh, eds, *African American English: Structure, History, and Use.* London: Routledge.

Carver, Craig. 1987. *American Regional Dialects*. Ann Arbor: U. of Michigan Press.

Cassidy, Frederic G., and Joan Hall 1985-. *Dictionary of American Regional English.* Cambridge: Belknap/Harvard U. Press.

Gilliéron, Jules. 1902--10. *Atlas linguistique de la France*. Paris: Champion.

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

Kirk, John, and William A. Kretzschmar, Jr. 1992. Interactive Linguistic Mapping of Dialect Features. *Literary and Linguistic Computing* 7.168--75.

Kretzschmar, William A., Jr. 1992. Isoglosses and Predictive Modeling. *American Speech* 67:227--49.

----- 1996a. Foundations of American English. In *Focus on the USA,* edited by Edgar Schneider, 25-50. Philadelphia: John Benjamins.

----- 1996b. Dimensions of Variation in American English Vocabulary. *English World-Wide* 17.189-211.

----- 1996c. Quantitative Areal Analysis of Dialect Features. *Language Variation and Change* 8.13-39.

----- 2003a. Linguistic Atlases of the US and Canada. In *Needed Research in American Dialects*, ed. by Dennis Preston. *Pubs of the American Dialect Society* 88: 25-48.

----- 2003b. Mapping Southern English. *American Speech* 78: 130-149.

----- 2004. Southern English by the Numbers. LAVIS III, Tuscaloosa, 2004.

----- 2009. *The Linguistics of Speech*. Cambridge: Cambridge University Press.

----- In Press. Evidence about Profiling from Linguistic Survey Research. In John Baugh, ed., *Linguistic Profiling in Global Perspective* (Ford Foundation).

-----, Jean Anderson, Joan Beal, Karen Corrigan, Lisa Lena Opas-Hanninen, and Bartek Plichta. 2006. Collaboration on Corpora for Regional and Social Analysis. *Journal of English Linguistics* 34: 172-205.

-----, and Rafal Konopka. 1996. Management of Linguistic Databases. *Journal of English Linguistics* 24: 61-70.

-----, and Deanna Light. 1996. Mapping with Numbers. *Journal of English Linguistics* 24.343-57.

----- and Edgar Schneider. 1996. *Introduction to Quantitative Analysis of Linguistic Survey Data: An Atlas by the Numbers.* Thousand Oaks, CA: Sage.

----- and Susan Tamasi. 2003. Distributional Foundations for a Theory of Language Change. *World Englishes* 22: 377-401.

-----, et al. 1993. *Handbook of the Linguistic Atlas of the Middle and South Atlantic States.* U. of Chicago Press.

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

------ 1972. *Studies in Area Linguistics*. Bloomington: U. of Indiana Press.

------ and Raven I. McDavid, Jr. 1961. *The Pronunciation of English in the Atlantic States*. Ann Arbor: U. of Michigan Press.

Kurath, Hans, et al. 1939. *Handbook of the Linguistic Geography of New England.* Providence: Brown U. for ACLS.

----- 1939-43. *Linguistic Atlas of New England.* Providence: Brown U. for ACLS.

Labov, William. 1963. The Social Motivation of a Sound Change. *Word* 19:273-309.

------ 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.

----- 1991. The Three Dialects of English. In P. Eckert, *New Ways of Analyzing Sound Change* (Orlando: Academic), 1-44.

------ 1994. *Principles of Linguistic Change*. Oxford: Blackwell.

Labov, William, Charles Boberg, and Sherry Ash. 2006. *Atlas of North American English: Phonetics, Phonology and Sound Change.* Berlin: Mouton de Gruyter.

Lee, Jay, and William A. Kretzschmar, Jr. 1993. Spatial Analysis of Linguistic Data with GIS Functions. *International Journal of Geographical Information Systems* 7: 541-560.

McDavid, Raven I., Jr., William A. Kretzschmar, Jr., et al. 1982--86. Basic Materials: Linguistic Atlas of the Middle and South Atlantic States and Affiliated Projects.

Chicago Microfilm MSS on Cultural Anthropology 67.360-75. Chicago: Regenstein Library Photoduplication Dept., U.of Chicago. (microfilm).

McDavid, Raven I., Jr., and R. O'Cain. 1980. Linguistic Atlas of the Middle and South Atlantic States. Fasc. 1-2. Chicago: U. of Chicago Press.

Montgomery, Michael. 1998. The Treasury of LAGS: Its History, Organization, and Accomplishments. In *From the Gulf States and Beyond: The Legacy of Lee Pederson and LAGS*, edited by M. Montgomery and T. Nunnally, 167-85. Tuscaloosa: University of Alabama Press.

Nerbonne, John. 2004. Aggregate Variation in the South in LAMSAS. LAVIS III, Tuscaloosa.

Pederson, Lee. 1981. *Linguistic Atlas of the Gulf States: Basic Materials.* Ann Arbor: UMI.

----- 1986. A Graphic Plotter Grid. *Journal of English Linguistics* 19:25--41.

----- 1987. An Automatic Book Code (ABC). *Journal of English Linguistics* 20:48-71.

----- 1988. Electronic Matrix Maps. *Journal of English Linguistics* 21:149--74.

----- 1986-92. *Linguistic Atlas of the Gulf States.* Athens: U. of Georgia Press.

----- 1993. An Approach to Linguistic Geography: The Linguistic Atlas of the Gulf States. In *American Dialect Research*, edited by D. Preston, 31-92. Philadelphia: John Benjamins.

Thill, J., W. Kretzschmar, I. Casas, and X. Yao. 2008. Detecting Geographic and Socio-economic Associations in English Dialect Features with Self-Organising Maps. In *Self-Organising Maps: Applications in GI Science*, edited by P. Agarwal and A. Skupin, 87-106. London: Wiley.