# X-RAY MICROBEAM
# SPEECH PRODUCTION DATABASE
# USER'S HANDBOOK

### Version 1.0 (June 1994)



JW12     "[uei]..[au]"
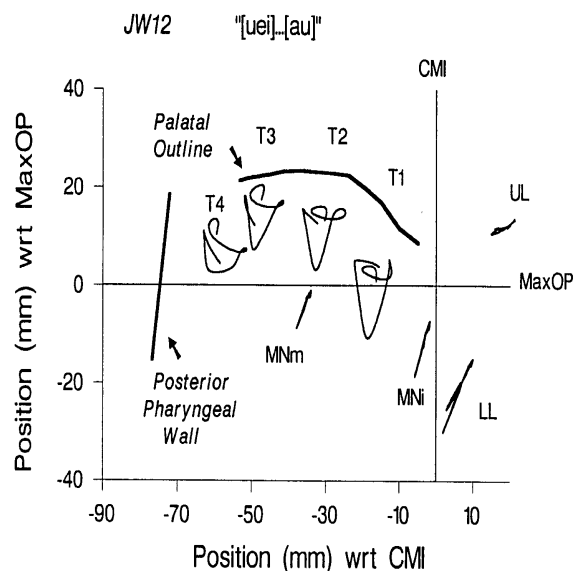
**prepared by**
## John R. Westbury

**with the assistance of**
**Greg Turner & Jim Dembowski**

**Waisman Center on Mental Retardation & Human Development**
**University of Wisconsin**
**Madison, WI 53705-2280**

**About the cover...**

Well, yes Virginia, there is a vowel triangle.
Fleshpoints along the tongue
trace this triangle as a speaker repeats vowel sequences
like "[uei]...[au]."  How interesting!
Perhaps for the curious reader,
it is also interesting to note that fleshpoints
on the lips and lower jaw do not trace triangles
during the same task.

# TABLE OF CONTENTS

# FOREWORD

A while back, I was idly thumbing my *Thesaurus of Words and Phrases*, and was struck dumb, as they say, by a rush of "deja vu all over again," sparked by Roget's Preface to his First Edition (1852):

> *It is now nearly fifty years since I first projected a system of verbal classification similar to that on which the present Work is founded. Conceiving that such a compilation might help to supply my own deficiencies, I had, in the year 1805, completed a classed catalogue of words on a small scale....Believing that a repertory of which I had myself experienced the advantage might, when amplified, prove useful to others, I resolved to embark in an undertaking which, for the last three or four years, has given me incessant occupation, and has, indeed, imposed upon me an amount of labour very much greater than I had anticipated. Notwithstanding all the pains I have bestowed on its execution, I am fully aware of its numerous deficiencies and imperfections, and of its falling far short of the degree of excellence that might be attained. But, in a Work of this nature, where perfection is placed at so great a distance, I have thought it best to limit my ambition to that moderate share of merit which it may claim in its present form; trusting to the indulgence of those for whose benefit it is intended, and to the candor of critics who, while they find it easy to detect faults, can at the same time duly appreciate difficulties.*

A fine first defense, for his project, or ours.

I have often viewed a reasonable goal of this project to be a rather long analogue of Perkell's (1969) *Physiology of Speech Production*, the now-classic cineradiographic account of thirteen disyllables spoken by a single male talker. On that thin volume, those of my generation cut their speech kinematic teeth. In the database described on the following pages, there are data from more than fifty talkers, each represented by almost eighteen minutes of speech. It has taken much more work from members of our group than I first imagined, in no small part because our intent, like Roget's, has been to produce a public resource that might benefit many, for years to come. Our physiological view of each utterance is coarser than Perkell's images, since ours are microbeam data: motions of points in a plane. However, our views across speakers and tasks are many times richer. We hope these will be of use to those who take the time.

Many have contributed directly to this project: Todd Brennan, Jim Dembowski, Nancy Gasper, Qiang Guo, Michiko Hashi, Carl Johnson, Ray Kent, Rick Konopacki, Mary Lindstrom, Paul Milenkovic, Jim Myers, Chuck O'Hare, Janine Tonsoni, Greg Turner, Gary Weismer, Quizhen Xue. Others have contributed significantly, but less directly. Accordingly, it is sometimes hard to give full credit, or lay proper blame. As far as we know now, we are not yet done. This is merely today's pickle, and we are in it.

. . . . . . . . . . . . . . . . . . . . . . . . J.R.W.

# CHAPTER ONE
## An Introduction to the X-ray Microbeam Speech Production Database

In 1989, we proposed (to NIDCD) the idea of developing a large-sample speech production database, incorporating point-parameterized representations of lingual, labial, and mandibular movements, recorded from the University of Wisconsin (UW) X-ray microbeam (XRMB) system in synchrony with the resulting acoustic sound pressure wave. The database was intended to span a relatively large number of speakers, and a rich, uniform inventory of utterances and oral motor tasks, all described according to a common kinematic framework. The resulting resource was intended to be sufficiently accurate and deep to withstand statistical scrutiny of variance, within and across speakers, and perhaps most importantly, to be an **open resource**, available for unlimited inspection and use by other speech scientists. The dataset described in this handbook represents the product of what seems even now a long-lived series of acquisition and processing efforts. The intent of these efforts has been to make the data streams and supplementary materials as accessible, and as free of error, as possible.

*Rationale*

The rationale for developing such a resource is multi-faceted, but also easy to understand. There is intrinsic value in a normative, statistical description of speech movements, against which disordered articulatory behavior can be compared. The requirement for descriptive, normative databases is readily accepted in health-delivery fields, and enormously detailed investigations, sometimes spanning many years, but always spanning large subject samples, have been devoted to the problem. For example, cephalometric (e.g., Broadbent et al., 1975; Behrents, 1985, a-b) and anthropomorphic (Farkas, 1981) studies documenting life-span changes in the craniofacial skeleton and overlying soft tissue have proven to be essential resources for orthodontic management of malocclusions, and surgical reconstruction of congenital and acquired skeletofacial malformations. Treatment objectives, in either case, always entail both the approximation and/or maintenance of normal function, and a normal, pleasing appearance. "Knowing the normal" (LeSavoy, 1981, p.2), through lengthy, detailed study and "extensive, structured experience" (Rosenbek & LaPointe, 1985, p. 112), must therefore precede diagnosis and treatment in any clinical discipline. Therapeutic intervention presupposes a valid, statistically-defensible sense of normalcy, quantitatively stated, and subject to inspection by, and disclosure to, others in the clinical community.

A second argument for large databases emphasizes the value such resources have relative to **informed** hypotheses that must underlie a general theory of speech production. The description of systematic variation in speech performance is a primary object of explanation for theory. Naturally, we cannot explain before we describe. Moreover, we cannot base a general theory, explaining why speech motor and perceptual behaviors vary as they do, on phenomena that are only narrowly described. Instead, broad, detailed descriptions of speech perception and production provide the necessary knowledge base for uncovering general regularities in behavior, for prompting ideas to account for those regularities, and for testing hypotheses arising from

other, smaller data sets or independent lines of thought (Maddieson, 1984). This type of argument justifies the compilation of cross-language databases such as the Stanford Phonology Archive, the UCLA Phonological Segment Inventory Database (UPSID), and the Child Language Data Exchange System (ChilDES).

A third argument for developing shared databases underscores the productive benefit of community resources, and the potential economy of such efforts. Maddieson (1984, p.3) has pointed out, for example, that "once a [representative, extensive, and uniform] database is established, numerous commensurate studies on the same data can be made." This claim is supported by frequent citations of the UPSID (cf., the collection of papers edited by Ohala, 1989); common sharing and continuing analyses of articulatory data obtained from the University of Tokyo x-ray microbeam system (cf., collected papers and citations therein, edited by Fujimura, 1988); and, by continued reliance on large-scale statistical characterizations of English (e.g., Hultzen et al., 1964; Kucera & Francis, 1967; Carterette & Jones, 1974). These statistical characterizations have themselves served as definitive source materials for more recent analyses of large data sets, and development of dense utterance inventories, that have been required to formulate and evaluate speech-synthesis and recognition schemes (cf., Umeda, 1975, 1977; Crystal & House, 1982, 1988a-d; Fisher et al., 1986; Lamel et al., 1986; Klatt, 1987; Elman & Zipser, 1988).

The need for large data sets relating to speech production, encompassing many speakers and tasks, has been frequently noted (cf., Abbs, 1986; Klatt, 1986; Perkell, 1986; Pisoni, 1986), and is relatively easy to satisfy if the primary interest is with certain acoustic consequences of speech. The sound pressure wave is easy to record. However, prior to the development of the XRMB system (Fujimura, Kiritani, & Ishida, 1973; Kiritani, 1986) and equivalent techniques (Perkell, 1988), it was not technologically feasible to collect and analyze large, standardized datasets describing normal speech motor performance, particularly when the desired information included kinematic data for the tongue. Reviews of contemporary publications describing tongue movements, for example, show that speaker samples of such studies are generally small (e.g., three or fewer speakers in more than 80% of papers), and that research objectives, speech samples, and analysis conventions vary considerably.

Small-sample physiological datasets are tractable, and can be indispensable for motivating new theory, and evaluating certain "hard" hypotheses. At the same time, however, small samples are problematic in that it is impossible to infer from them, with any conventionally acceptable level of confidence, the true underlying distribution of behavioral patterns. Pooling results across studies can circumvent this problem, but only when the recording conditions, data types, and representation and analysis conventions are sufficiently similar from one study to the next. Careful readings of the literature show that this is rarely the case (Westbury, 1994). Consequently, a serious problem with small sets, even those that are narrowly-focused, is their lack of generality. Lubker & Gay (1982, p. 442) explicitly acknowledged this limitation some years ago, writing that "the use of more subjects is not just 'suggested' [in studies of speech production], we believe that it is mandatory."

A large-sample production database should speak directly to this requirement. An additional, attractive feature of such a database is that it should yield data suitable for examining many issues that have occupied the attention of speech researchers. At the time of our original proposal, we suggested a number of specific questions that could easily be addressed using materials like those we have collected. Over the last couple of years, we have begun to analyze database materials with those questions in mind. But, it is also true that many additional, non-trivial questions can, and *should*, be addressed using the same materials. The need for substantial additional articulatory research, such as these materials make possible, has been explicitly stated in the NIDCD (1989) *Report of the Task Force on the National Strategic Research Plan*: "First, information is needed about the temporal sequencing of the articulators for the production of a given phonological message, and how this temporal sequencing changes with changes in stress or rate of speaking. Second, information, constrained by equations of motion, is needed about the shapes that articulatory trajectories can assume. Third, a model of the articulators in the vocal tract, which can be used to synthesize speech, is needed" (p. 139).

The UW XRMB facility has afforded a unique setting within which a large-scale investigation of speech movement is possible. The facility has allowed us to record and visualize many of the critical movements of speech, in conjunction with synchronous physiologic data and at relatively low risk to speakers. These records provide a detailed representation of speech production kinematics in the context of an analysis environment optimized for efficient data reduction. It is sobering to recall that the general objectives for such an investigation are now no different from those formulated for high-speed cineradiography more than 30 years ago by Öhman and Stevens (1963, p. 1889), and include:

> *(1) obtaining further understanding of the relationships between the positions of the articulatory structures, vocal-tract configurations, and acoustic outputs; (2)examining the dynamic properties of the various components of the articulatory mechanism and the interactions among the movements of these components; and (3) gaining an understanding of the transformation from a discrete linguistic description of speech to the continuous motions of the articulatory structures.*

What is still required to address each of these broad objectives, above all, are accurate data, uniformly described, representing a wide range of performances by many performers.

# CHAPTER TWO
## A Short History of the UW XRMB Facility

Those of us who are here now weren't here from the beginning. For that reason, it is hard to piece together a complete view of what happened, when, and why, over the years when the facility was developed and finally put to use. Up to a point, we have to rely on surviving paper and the memories of others to figure these things out. Our own memories take over after a while. These qualifiers, probably the standard lament of historians, amount to an apology for inaccuracies in the account that follows.

The XRMB system is a device for recording motions of articulators during speech production. It performs this task by generating a very narrow beam of high-energy x-rays, and rapidly directing this beam, under high-speed computer control, to track the motions of 2-3 mm diameter gold pellets glued to the tongue, jaw, lips, and soft palate. The UW system is a second-generation device, patterned in part upon an earlier system developed by JEOL of Tokyo, Japan, for the Research Institute of Logopedics and Phoniatrics at the University of Tokyo.

The idea that a second-generation system could and should be built must be credited to Osamu Fujimura, and its existence must be seen as part of his legacy to the field. Certainly, Fujimura can be recognized as the inventor of the microbeam technique, and was central to the development of the earlier Tokyo device. A number of different institutions were informally approached as potential hosts for the new system. For a variety of reasons, UW became the choice, and the idea that the new system should be there must have been firmly in place by 1978. We can guess this since the initial grant proposal to NIH (J.H. Abbs, P.I.) was submitted in October, 1979. Funding to support system design, construction, and eventual use, was first received on September 1, 1980. The system was to be fabricated and initially tested by a group of physicists and engineers at UW Physical Sciences Laboratories (PSL: Murray Thompson, director), and was designed to improve upon the earlier Tokyo system in a number of ways. One of the more significant of these was the intent that the UW system promote "efficient use for the acquisition of large data bases with the system being shared by many speech researchers in the [United States]" (Thompson, 1984, p.1). From the beginning, the UW XRMB system was promoted to both the funding agency and speech science community as the central component of a nationally-shared laboratory facility that would permit "simultaneous and real time sampling and storage of wide-band speech acoustics, EMG from multiple speech muscles, and associated aerodynamic events [...in the context of] state-of-the-art computer processing, display, and modeling algorithms[...] to optimize the meaningful interpretation of these multiple signals" (Abbs & Nadler, 1987, p. 14). An important rationale for the development of the UW system was that "the multiple electrophysiological, biomechanical, aerodynamic, and acoustic processes of human speech production require the capability for transduction and analysis of large quantities of multi-faceted physiological data" (Abbs & Nadler, 1987, p. 14).

An early resource estimate from PSL suggested that system design and fabrication, to first use, would require something on the order of 14-15 man years. This estimate encompassed

development of a theory of operation, physical construction of the device's massive frame and shell, fabrication and/or acquisition of many key components (e.g., high voltage power supply [HVPS], isolation transformer, scintillation counter, cooling systems, and ion pumps), acquisition of computing hardware, and development of computer code necessary to initiate and terminate beam control and tracking. The resource estimate proved reasonably accurate, judging by the fact that first operation of the system, involving generation of the high-energy X-ray beam, occurred in January, 1984. Many problems seemed to plague the system around this time, having to do with stability of the HVPS, vacuum maintenance along the electron-beam line, filament longevity, leaks and impurities in the sulfur hexafluoride "bath" surrounding the capacitor stack, and recurring damage to the tungsten target. At this time, the system in all its large and weighty glory (even then at some 10-12 tons of lead, with dual 5-meter components: the beam line and capacitor stack), was still in place at its fabrication site (Stoughton, WI, well off the UW-Madison campus). It would not be moved to its subsequent home laboratory, at the far west end of the Madison campus, for at least another 18 months. A second-cycle five-year grant proposal to NIH, for ongoing development and eventual use, was submitted in October, 1984. Shortly thereafter, first use of the system for generating full-field scans, and for real-time tracking [4 (lead) pellets on a phantom], seems to have occurred on the first of November, 1984, roughly four clock years after first funding.

The system was transported to the UW Waisman Center on Mental Retardation and Human Development, and reassembly was begun, about the first of October, 1985. For the next year and five months, it is likely that a great deal of software development occurred, in anticipation of use of the system in "live" experiments. The first such experiment with a human subject occurred in February, 1987, some six and one-half years after first funding. During the 1987 calendar year of system operation, approximately 30 separate experimental sessions were run. During calendar 1988, 81 new sessions were conducted; and during the first nine months of 1989, another 62 were added. In October, 1989, a third-cycle five-year grant proposal was submitted to NIH, and five months later, after 33 additional experiments, system use was halted for a significant upgrade in shielding, in anticipation of increasing the system accelerating voltage from 450keV to 600keV, producing "harder" photons and extending the use of the device for tracking pellets in the context of dental fillings. Prior to this time, all subjects run on the system had essentially been filling-free. System use resumed shortly before Christmastime, 1990, and continued under NIH support through the end of October, 1991. In that period, 72 additional experiments were conducted, 41 of which are represented as speakers in the X-ray Microbeam Speech Production Database.

Our proposal to develop the Speech Production Database had been submitted to NIH late in 1989, in parallel with, but independently of the core facility grant proposal, and first funding was received in July, 1990. It was our conviction at that time, and remains so now, that the UW XRMB facility was uniquely well-suited to the development of an open, large-scale database. In this sense, the Database, whose projected future value we believe to be both great and lasting, is itself testimony to the wisdom of having built the facility.

During the most active phase of data acquisition, spanning roughly 2.5 years from late

1987 to early 1990, and a subsequent 10-month period spanning December, 1990 to October, 1991, access to the system for all experimentation, by all on-site and off-site scientists, was governed by procedures set by a User's Committee, a group of eight or so off-site speech scientists not directly affiliated with the facility. This committee was charged broadly with the responsibility of providing "advice regarding directions of the [XRMB facility] in establishing and modifying its organization... and for overseeing the performance of the Administrative/ Technical Board and executive officers... to insure that the [facility served] its purpose" (Abbs, 1989, p. 125). Early members of the committee included Joe Perkell, Kathy Harris, Peter Ladefoged, & Ray Daniloff. At its last meetings, the committee was chaired by Maureen Stone, and included as members Cathe Browman, Bjorn Lindblom, Kevin Munhall, Sandra Hamlet, Jeri Logemann, Perkell, and John Westbury.

The reviewers' response to the third-cycle (1989) NIH proposal was insufficiently enthusiastic for continued funding of the facility-core grant, and for that reason, one year of transitional support was initiated in October, 1990, with the understanding that a revised proposal would be submitted, as it was in February, 1991. Unfortunately, that proposal was received with no more enthusiasm than the draft of a year earlier, and NIH support for the core facility was halted at the end of October, 1991. This was a difficult and anxious time for the Database project, which had been independently funded to acquire data from a system whose ongoing existence then seemed in great jeopardy.

Fortunately, after October, 1991, facility operation was maintained, though at a more modest level than in past years, with core support for its technical staff supplied by a series of collaborative research contracts, first between UW and ATR Auditory & Visual Perception Research Laboratories, and subsequently between UW & ATR Human Information Processing Research Laboratories (Kyoto, Japan). These contracts provided system access to ATR scientists interested in speech production, and to members of the UW research team responsible for the Speech Production Database. During the 2+ years spanning these contracts, some 66 new experiments have been conducted on the microbeam system.

Looking at the ever-popular bottom line, something on the order of eight million (direct) dollars were spent on development, maintenance, and operation of the system over a span of 13+ years. A large proportion of these funds were spent before any experiment was ever run. Costs over the last few years have been markedly lower than in previous years, where the initial investment was necessarily high. In relative terms, a large body of data has been collected on the system, spanning some 330+ experimental sessions, conducted by or on behalf of some 40-odd researchers or research teams, using more than 200 different speakers, and encompassing more than 3600 minutes of tracking time. These few statistics are deceptively simple, and to an unpracticed eye, may fail to reveal significant benefits of the facility. One such benefit is that the large aggregate collection of data obtained at the facility is essentially uniform in quality, and in certain representational conventions that significantly impact analysis and interpretation.

From a socio-historical point of view, if there is such a thing, the XRMB facility might be thought of as yet another speech science experiment: an attempt to determine whether a

complex, expensive shared facility could advance the field in ways that less complex, less expensive, more distributed facilities could not. This issue, with its associated and frequently fervent dialogue, is familiar to other branches of science, and to much of modern life, where citizens of varied communities routinely face the question of resource-sharing whenever elaborate resources of interest (e.g., libraries, computer mainframes, particle accelerators, swimming pools, schools) are beyond the means of individuals interested in them. It is common in such circumstances to pool resources for such facilities, and to forge contracts governing their access and use. The survival of shared resources then depends sometimes upon a collective sense of their intrinsic worth and fair administration, other times upon measurable estimates of benefit derived from them, and still other times upon continuation of the set of circumstances that gave rise to initial cost-sharing arguments.

In physical sciences such as physics and astronomy, and more recently, biological sciences such as cytogenics, the debate between big (usually shared-resource) science, and little science is familiar. For Speech Science, the XRMB facility was big science. There are many who believe big science initiatives to be intrinsically problematic, and therefore to be routinely avoided, dismissed out of hand as bad policy, perhaps because by its nature big science robs too many of an opportunity to do science "outside the magic circle." The fact that big science sometimes wins out, in spite of this general bias against it, results from the extraordinary expense and operating conditions associated with some facilities (e.g., astronomical observatories) required to do certain types of work at all. Small facilities that provide equivalent function just cannot be built. Thus, it is sometimes true that big, shared facilities are simply the cost of doing respectable business. An important, open question in speech circles at this time, is whether microbeam-equivalent data can be obtained by other methods. The probable answer seems to be yes. If so, perhaps we can rationalize the eventual loss of the facility as a shared resource. However, this rationalization reduces the idea behind the facility to the belief that it represents nothing more than a data type, dispensable once similar data can be otherwise acquired. This unfortunate view misses the mark badly. The facility was never merely a data type. Rather, it was to have been a shared laboratory, capable of processing many data types, and capable, given support, of providing new data types as times and ideas change. Why that notion has run aground is a mystery.

First funding of the facility, from NIH, may have been due as much to the quality of the proposal, and the attraction of the idea, as to some well-placed lobbying by senior scientists in the field. At that time, there was apparently a strong consensus that the facility and its funding were a very good idea. There is no question that the facility was promoted as a nationally-shared research resource, funded virtually entirely by a single grant to UW, but open for use by all scientists with interests in the movements of speech, associated muscle actions, and their aerodynamic consequences. To that end, a sufficient array of general-purpose instruments, and expert staff of technicians, were recruited and employed to make complex, multi-channel experiments routinely feasible. Those who do such work know that it is much harder to do well, and intelligently, than many might imagine. As with all things, only time will tell whether the great cost and effort were worth it.

# CHAPTER THREE
## Speech and task sample

Speech research is nothing if there is nothing to say. In fact, many would argue that speech research is nothing if there isn't just the right thing to say. The usual wisdom is that production tasks must be surgically precise if they are to elicit behaviors of interest. If tasks of some desired type, cleanly counter-balanced by condition, do not readily occur in the speaker's native language, many in our field have no compunction about making them up: to wit, the ever-popular nonsense utterances that form a basis for much of what we think we know about the phenomena and principles of articulation.

The philosophy of the Speech Production Database was somewhat different. In general, task sequences were intentionally not balanced by condition. To the extent that anything is known well in speech science, discrete "segmental" phenomena from balanced protocols are reasonably well-understood. Instead, and by design, the Database task inventory sampled widely from relatively natural material. The hope was more that users would find some interesting materials, rather than precisely their favorite materials, for future analysis.

### 3.1. Rationale

A lengthy, uniform task inventory, representing several different task types, was intended for every speaker. This inventory represented a design compromise between two points of view. The first required that the task list be sufficiently broad to encompass most of the range of motor and linguistic tasks a speaker has to perform when talking. The second required that the list be sufficiently redundant to provide meaningful estimates of intra-speaker variability, and thereby allow reliable inferences regarding speaker intent, and control principles governing the speech act. We purposely chose a greater emphasis on breadth than redundancy, partly in deference to the project goal of developing a resource that can stimulate new ideas about speech production. The stimulation factor is plausibly greater for inventories containing many tasks and types, rather than a few tasks repeated many times. But, it is also important to understand that the microbeam technique does not favor highly redundant inventories. X-ray exposure times escalate rapidly with repetition; so too does the number of trials, unless repeated material is tightly packed. Exposure and trials need to be minimized if microbeam experiments are to be successful.

Three principal considerations governed the selection of tasks for the Database inventory: (1) historical precedent (e.g., stimuli known from other, seminal studies of speech production); (2) clinical application and relevance; and, (3) representativeness (e.g., phone and phone-sequence frequencies, speaking modes, suprasegmental manipulations). The task inventory reflects a heavier weighting toward more natural, connected speech than is customary in production experiments. There are two reasons for this. First, descriptions of connected speech, where speakers must contend with the vagaries of syntax, suprasegmentals, meaning, and uncommon sound sequences occurring across word boundaries, are generally less developed than those of citation speech. The few existing reports regarding the former (e.g., Umeda, 1975,

1977; Crystal & House, 1982, 1988a,b,c,d) suggest that speaking mode has a profound effect on many of the accepted segmental regularities that have been established for nonsense words and phrases, in citation-form speech.  A second reason for including a rich connected-speech component is that clinical assessments of impairment frequently depend upon subjective evaluation of speech samples that are at least similar in kind (cf. Templin & Darley, 1960; McDonald, 1964; Darley et al., 1975; Yorkston & Beukelman, 1981).

*3.2. Task list*

Tasks are listed below.  Percentages associated with task-type sub-headings indicate the approximate aggregate tracking time per type.  All items from the inventory that were repeated more than once in the experimental task list are followed by *(xN)*, where N represents the total number of repetitions.  The organization of tasks, in the utterance list, was the same for all speakers, and the position of each task (by *record*: see below) within the utterance list is indicated in curly brackets.  By design, tasks of different types were sprinkled throughout the utterance list, in the hope that mixing tasks would offset fatigue and drudgery, for speakers and experimenters.  A few tasks believed to be mechanically risky (e.g., prone to break adhesive bonds of pellets: *swallows*; *diadochokinesis*; and *maximal tongue/lip protrusion*), were restricted to occur late in the inventory.  The bracketed capitol letter [P] indicates that the task was also used in a practice session preceding attachment of pellets (see section 3.3 below).

Connected Speech: Prose passages (13%)

"Grandfather Passage" (Darley, Aronson & Brown, 1975, p. 298):  [P: first six sentences] {11-12}

> *You wish to know all about my grandfather.  Well, he is nearly 93 years old, yet he still thinks as swiftly as ever.  He dresses himself in an old black frock coat, usually several buttons missing.  A long beard clings to his chin, giving those who observe him a pronounced feeling of the utmost respect.  When he speaks, his voice is just a bit cracked and quivers a bit.  Twice each day he plays skillfully and with zest upon a small organ.  Except in the winter when the snow or ice prevents, he slowly takes a short walk in the open air each day.  We have often urged him to walk more and smoke less, but he always answers, 'Banana oil!'  Grandfather likes to be modern in his language.*

"The Hunter Script" (Crystal & House, 1982, p. 715):  [P: first five sentences] {78-81}

> *In late fall and early spring the short rays of the sun call a true son of the out-of-doors back to the places of his childhood.  Tom Brooks was such a man.  Each year at these times his desk seemed like a stone whose weight made him wish for the life he knew as a boy.  In the five years since leaving college he had not revisited his old haunts before.  But this March Tom found himself by a small stream with a gun at rest in the crook of his arm.  The desk that had tied him down was gone and his one thought was for quail.  He had been on the trail since dawn, but not one bird had crossed his path.  It seemed as*

9

*though five years without hunting had made him lose touch with all the small signs that he once knew -- signs that would tell for sure if an animal was near or not. Once he thought he saw a bird, but it was just a large leaf that had failed to drop to the ground during the winter. Tom stopped near a stream to rest. Soon after he had laid down his gun, he heard the sound of wings from across the stream, and five large birds came out of the brush. They flew to the edge of the stream unaware of the hunter. Tom placed his hand on his gun quietly. Slowly he raised it to his shoulder and took aim. The seconds ticked off like hours, but still the birds drank. Quick shots rang out. The years of waiting seemed to disappear with the successful culmination of the hunt.*

Counting, and sequences of number names (6%)

Number names [P] {88}

| | | | |
|---|---|---|---|
| one | two | three | four |
| five | six | seven | eight |
| nine | ten | eleven | twelve |
| thirteen | fourteen | fifteen | sixteen |
| seventeen | eighteen | nineteen | twenty |

Phrases made from number-name sequences

| | | |
|---|---|---|
| 9739286 [P] {3.1} | 8495571 [P] {3.2} | 5945341 [P] {3.3} |
| 4375125 {92.1} | 3647962 {92.2} | 1146327 {92.3} |
| 6582269 {63.1} | 7217424 {63.2} | 2315483 {63.3} |
| 7789388 {51.1} | 8761335 {51.2} | 2918524 {51.3} |
| 5681998 {72.1} | 6744166 {72.2} | |

 Oral Motor Tasks (8%)

A. Replicative jaw "wagging" {106}, and sequences of [sa] [P] {105}
B. Maximal tongue and lip protrusion {117,118}
C. Diadochokinesis -- puh, tuh, kuh  [P: puh only] {102-104} (A fourth DDK task,
          puhtuhkuh, was added for last five speakers recorded.)
D. Swallowing -- 2cc and 10cc (x5, each) {107-116}

Citation Words, near-words, sounds and sound-sequences (From Kucera & Francis (1967); and Weismer et al. (1988); Rodman, Moody & Price (1984); modelled on Peterson & Barney, 1952) (33%)

Citation Words:

about (x5) {18.3,18.6,41.3,76.3,95.4}          across (x5) {27.3,28.4,37.4,47.5,90.7}
above {54.5}                                    against {62.1}

almost {8.4}
already {6.6}
although {4.6}
among {47.2}
around {58.1}
back (x5) {8.3,18.7,76.4,100.2,100.5}
beautiful {33.1}
because {73.7}
become {44.4}
been {9.7}
before {32.2}
began {65.2}
between {28.5}
blend (x5) {6.7,23.2,32.4,65.4,91.6}
blink {95.2}
both (x5) {21.6,58.2,58.7,66.3,90.6}
but (x5) {4.3,22.7,41.1,90.1,95.3}
cash (x5) {25.3,65.3,66.5,73.1,89.2}
child {23.4}
children (x5) [P] {1.2,2.6,22.3,25.5,28.6}
coat (x5) {23.1,49.7,65.1,95.6,99.3}
conversation {89.4}
could {27.7}
country (x5) {21.1,21.7,49.4,58.3,66.2}
didn't {52.5}
dorm (x5) {25.1,32.1,33.4,49.6,52.2}
dormer (x5) [P] {1.3,6.4,22.2,23.5}
dormitory (x5) [P] {1.5,18.1,47.1,83.6}
early {41.6}
enjoy {61.3}
even {2.4}
find {47.6}
first (x5) {18.5,62.3,76.2,87.1,99.6}
flip {65.7}
form {8.1}
from {62.5}
glowing {4.7}
going {25.6}
had (x5) {22.6,23.6,58.4,61.7,89.7}
hail (x5) {6.1,21.5,83.4,83.7,91.1}
has [P] {1.7}
have {35.5}
head {87.3}
himself (x5) {83.2,89.1,90.5,95.1,100.1}

house (x5) {22.5,37.3,41.5,52.6,70.6}
information {54.4}
light (x5) {18.4,54.2,70.4,91.2,91.4}
long (x5) {37.2,54.1,61.4,62.4,87.4}
look {5.3}
major {76.1}
making {25.7}
man {61.2}
measure {95.5}
moment (x5) {5.4,5.6,23.7,28.7,44.3}
much (x5) {35.1,61.5,70.2,90.3,99.7}
never [P] {1.4}
nothing (x5) {2.1,9.6,73.2,76.6,100.4}
order (x5) {9.1,35.2,49.2,52.4,99.2}
people (x5) {5.1,32.7,35.4,62.2,62.6}
point (x5) {41.2,61.1,73.3,83.3,89.3}
problem (x5) [P] {1.1,37.6,44.2,49.3,99.5}
program (x5) {28.1,52.1,52.7,66.6,90.4}
programmer (x5) {5.5,8.7,32.3,37.5,54.3}
pushed {65.6}
question {37.7}
quite {5.7}
right (x5) {6.3,9.5,27.4,87.7,99.4}
row (x5) {4.1,9.2,33.2,47.7,49.1}
school (x5) [P] {1.6,8.6,22.1,73.5,91.7}
second (x5) {27.6,33.6,47.3,66.7,87.2}
seemed (x5) {22.4,27.1,44.1,44.6,83.1}
sense (x5) {21.4,27.5,32.5,33.5,37.1}
ship (x5) {2.7,8.2,41.4,58.5,91.5}
shoot (x5) {9.3,61.6,66.1,66.4,89.6}
sigh {54.6}
silk {21.3}
sip {54.7}
smooth {35.3}
so {70.1}
special (x5) {2.5,4.2,4.4,35.7,70.7}
street (x5) {2.3,23.3,47.4,100.3,100.7}
than {33.3}
that (x5) {25.2,70.3,73.6,87.6,89.5}
there {70.5}
things (x5) {4.5,8.5,33.7,41.7,100.6}
this (x5) {2.2,6.2,28.3,65.5,76.5}
through {25.4}
told (x5) {5.2,6.5,28.2,32.6,49.5}

understand {21.2}
used {9.4}
wax (x5) {76.7,83.5,91.3,95.7,99.1}
weigh {62.7}
were {44.5}
what {73.4}
words {18.2}
work {44.7}
would {35.6}
yet {58.6}
yourself {27.2}
zero {90.2}


Citation sVd's  [P] {13}

| | | | | |
|---|---|---|---|---|
| [sid] | [sɪd] | [sed] | [sɛd] | [sæd] |
| [sʌd] | [sɑd] | [sɔd] | [sod] | [sʊd] |
| [sud] | [sɝd] | [sɔid] | [sɑid] | [sɑʊd] |

Isolated vowels  [P] {14}

| | | | | | |
|---|---|---|---|---|---|
| [i] (beet) | [ɪ] (bit) | [e] (bait) | [ɛ] (bet) | [æ] (bat) | [ʌ] (but) |
| [ɑ] (hot) | [ɔ] (bought) | [o] (boat) | [ʊ] (foot) | [u] (boot) | [ɝ] (dirt) |

Vowel sequences  [P] {15}

[iu]   [iɑ]   [uɑ]   [ɑu]   [ɑi]   [ui]

Citation VCV's  [P] {16}

| | | | | | |
|---|---|---|---|---|---|
| [ʌpɑ] | [ʌtɑ] | [ʌkɑ] | [ʌbɑ] | [ʌdɑ] | [ʌgɑ] |
| [ʌhɑ] | [ʌmɑ] | [ʌnɑ] | [ʌf/vɑ] | [ʌs/zɑ] | [ʌʃ/ʒɑ] |
| [ʌʧ/ʤɑ] | [ʌwɑ] | [ʌjɑ] | [ʌlɑ] | [ʌrɑ] | |

Sentences (40%)

DARPA/TIMIT Material:
She had your dark suit in greasy wash water all year.  (x5) [P,x3] {10.3,46.3,53.3,60.3,98.2}
Don't ask me to carry an oily rag like that.  (x5) [P,x3] {17.1,40.3,68.2,74.3,96.3}
That noise problem grows more annoying each day. {56.3}
They remained lifelong friends and companions. {36.1}
Shaving cream is a popular item on Halloween. {30.2}
She always jokes about too much garlic in his food. {45.2}

Grandmother outgrew her upbringing in petticoats. {77.2}
Does Creole cooking use curry? {96.1}
Combine all the ingredients in a large bowl.  (x20; rate & clarity) [P, slow & clear] {slow: 26,38,57} {fast:34,42,93} {clear:82} {normal:39.2,53.3,59.4,74.2,84.3}
The sermon emphasized the need for affirmative action. {56.1}
Hispanic costumes are quite colorful. {7.3}
Cheap stockings run the first time they're worn. {86.2}
A roll of wire lay near the wall. {96.2}
Don't do Charlie's dirty dishes. {10.2}
Elderly people are often excluded. {101.1}
I assume moisture will damage this ship's hull. {21.1}
The gorgeous butterfly ate a lot of nectar. {43.2}
The oasis was a mirage. (x2) {67.3,67.4}
Porcupines resemble sea urchins. {85.1}
When all else fails, use force.  (x5) {29.1,39.1,67.1,68.1,101.1}
We are open every Monday evening. {31.1}

Other sentence material:

Put these two back.  (x15; emphasis) [P] {emphasis:75,94} {normal (x6): 24.2,29.3,30.3,55.1,60.2,98.3}
Put this one right here. {86.1}
She is about two or three. {7.1}
It's just a little thing. {77.1}
Do they go up and down? {71.1}
They all know what I said.  (x5) {31.3,36.2,50.2,59.2,67.2}
When can we go home? [P,x3] {7.2}
I think that's real. {84.2}
The other one is too big. (x15; rate) {slow:26,38,57} {fast:34,42,93} {normal (x5): 10.1,36.3,45.1,50.1,97.1}
You can shoot at the ship or do nothing. (x5) {17.2,29.2,43.1,53.2,97.3}
The dormitory is between the house and the school. (x5) {19.3,40.2,74.1,86.3,101.3}
Second children are often special people. (x5) {24.3,48.2,56.2,64.3,85.3}
I'll make sense of the problem in a moment. (x5) {31.2,48.3,59.3,64.1,69.2}
The coat has a blend of both light and dark fibers. (x5) {19.1,20.2,55.3,59.1,69.3}
You must blend certain things to make a special wax. (x5) {17.3,40.1,43.3,71.2,83.1}
Things in a row provide a sense of order. (x5) {24.1,39.3,50.3,60.1,69.1}
Across the street stands a country school. (x5) {19.2,46.2,68.3,71.3,85.2,}
If I had that much cash I'd buy the house. (x5) {20.3,45.3,77.3,97.2,98.1}
The point of the program will be told before long. (x5) {30.1,46.1,48.1,55.2,64.2}

*3.3. Recording conventions*

Approximately nineteen (19) **tracking** minutes were required to record the full

inventory. For comparative purposes, a small subset of the task inventory (identified by [P], in the preceding list) was pre-recorded for each speaker, under identical conditions but without pellets. Pre-recording allowed speakers to become familiar with tasks and certain experimental conditions, prior to subsequent performances with pellets. Start-to-finish, approximately seven hours per speaker were necessary to obtain all data, including pre-recording; pellet-gluing; pellet-tracking; anthropomorphic measurement; dental model fabrication; and, hearing-screening and assessment interviews.

Not all tasks were recorded for all speakers. There are several reasons why this is so. In our two earliest sets (JW5 & 6), for example, only the *word, sentence,* and *swallow* tasks were recorded. At the time these speakers were recruited, we were still having system trouble recording and storing lengthy tasks (record lengths greater than 15 seconds). At the same time, system dosimetry standards were being revised, in conjunction with recent replacement of the XRMB system target. Accordingly, for those two speakers only short sets were attempted. For the majority of remaining speakers, full inventories were attempted, though a number of these will have unintentionally incomplete sets, due to various system failures in acquisition; detachment of key pellets during the experiment; or "accidents" (**not** involving radiation) of one odd sort or another. One speaker, for example, hadn't eaten dinner before we began a late-evening session. After about three hours, she decided she was so hungry she couldn't go on. Though we had already collected a substantial proportion of our target data from her, she was unable to return to complete the data collection. So, for her, we are irreparably short.

All tasks were organized and recorded in short subsets of material we refer to as *records*. Names given to records (e.g., *wordsNN, clearNN, hunter[a]*) reflect their general content. Records of the *word* type contained seven words selected randomly from the full list of citation words, each read with a brief pause in between; records of a *sentence type* contained three different examples selected quasi-randomly from the sentence list, also read with a short pause in between. The entire recording protocol consisted of a minimum of 118 records (excluding those repeated due to reading errors and/or acquisition flaws, and excluding those records holding initialization scans and certain calibration tasks). By type, records were distributed in the following way:

| Record type | N | Usual track time/rec (s) |
|---|---|---|
| word: standard | 40 | 7.5 |
| word-like: vcv, cvc, vseq, v | 4 | 10-27 |
| sentence: normal | 36 | 11 |
| sentence: fast & slow | 6 | 10 & 16 |
| sentence: clear | 1 | 17.5 |
| sentence: emphasis | 2 | 17.5 |
| paragraph | 6 | 22-25 |
| swallow: 2cc & 10cc | 10 | 3 |
| diadochokinetic | 3 | 3.5 |
| counting (1-20) | 1 | 20 |
| number sequences | 5 | 10 |
| oral gym: wag & protrude | 4 | 10 & 3.5 |

The order and content of records was the same for all speakers.  Record presentation during an experiment, by means of a video terminal placed at eye level and less than a meter in front of each speaker's face, was governed by a text file specifying record name, tracking time, task instructions, and tasks to be performed.  A full copy of the record list is reproduced in an appendix to this manual, partly to illustrate record order, but also to illustrate printed performance instructions accompanying each task.  Also in an appendix, is a count of the total number of records of each (task) type, obtained across all speakers in the Database.

During any microbeam recording session, the x-ray beam was "on" only during a fixed time window associated with each record, when the speaker was to perform the current task.  Speakers initiated task performance following an audible cue (an automatically delivered tone), only after the system located initial positions of all pellets by rapid local scans.  Once a record was initialized, the system tracked pellet motions for the duration of the record window.  At the end of the record, the X-ray beam was halted, and the acquired data transferred to archival storage.  Data transfer required roughly two times real time.  Consequently, sequential records in a record list were always separated by at least the amount of time required to flush memory, and re-initialize the system.  Additional adjustments (e.g., having to do with speaker position, fatigue, calls of nature, and the like) could make inter-record intervals relatively lengthy.

The length of each record, in tracking seconds, was governed by task content and the speaker's habitual reading rate, but was constrained above by a memory buffer limit on the sampling engine for the acoustic and accelerometer tracks (section 5.1.1).  The longest record duration, containing 21 citation vcv's of the form [əCɑ], was 27 seconds.  The record-length ceiling required partitioning long passages of connected speech (e.g., the Hunter and Grandfather passages).  For these partitioned passages, the last and first sentences of each contiguous subdivision were repeated, to avoid major discontinuities in reading rate and fluency.

Some tasks (e.g., those involving rate, clarity and emphasis manipulations) could not be practically woven amongst others, and were segregated into separate, homogeneous records.  A fast-speech record, for example, contained within a single record two different sentences, both to be read at "double" speed (i.e., twice the speaker's self-judged "normal" speaking rate), in alternating order.  Packing tasks of a common type within records, often involving some repetition, may suppress performance variability.

*3.4. Phonetic Content*

A modified DARPABET transcription of the inventory is reproduced in an appendix.  In this transcription, speech sounds are followed by white space, word-final sounds by asterisks, and sentence-final sounds by backslashes.  In transcription, the full inventory contains 6275 sounds, distributed according to phone frequency in the following way:

| Vowels: | | | Consonants: | | |
|---|---|---|---|---|---|
| phone | freq | %age | phone | freq | %age |
| AX | 430 | 6.85 | N | 407 | 6.49 |
| IH | 251 | 4.00 | S | 351 | 5.59 |
| IY | 226 | 3.60 | T | 331 | 5.27 |
| AH | 192 | 3.06 | R | 323 | 5.15 |
| AO | 165 | 2.63 | L | 300 | 4.78 |
| AE | 149 | 2.37 | D | 228 | 3.36 |
| EH | 148 | 2.36 | Q | 222 | 3.54 |
| OW | 128 | 2.04 | K | 217 | 3.46 |
| UX | 121 | 1.93 | M | 186 | 2.96 |
| AY | 119 | 1.90 | DH | 166 | 2.65 |
| AA | 89 | 1.42 | B | 153 | 2.44 |
| EY | 78 | 1.24 | Z | 134 | 2.14 |
| AXR | 65 | 1.04 | P | 124 | 1.98 |
| ER | 42 | 0.67 | W | 121 | 1.93 |
| AW | 35 | 0.56 | F | 114 | 1.82 |
| OY | 24 | 0.38 | HH | 103 | 1.64 |
| UH | 22 | 0.35 | V | 92 | 1.47 |
| | | | G | 71 | 1.13 |
| | | | SH | 67 | 1.07 |
| | | | NG | 62 | 0.99 |
| | | | TH | 62 | 0.99 |
| | | | DX | 51 | 0.81 |
| | | | Y | 41 | 0.65 |
| | | | CH | 34 | 0.54 |
| | | | JH | 27 | 0.43 |
| | | | H | 4 | 0.06 |

*3.5. Listening Assessment*

A coarse listening assessment of reading performance was completed for each speaker. Broadly, the assessment was intended to identify any of 20 word or phrase-level features per record per speaker, including the following: (1) phoneme deletion, (2) phoneme distortion, (3) phoneme substitution, (4) word substitution, (5) word addition, (6) word deletion, (7) word transposition, (8) inappropriate prosody, (9) hyperarticulation, (10) inappropriate pause, (11) sound-syllable repetition, (12) sound-syllable revision, (13) word repetition, (14) word-revision, (15) phrase repetition, (16) phrase revision, (17) truncation, (18) dysfluent reading, (19) phoneme addition, and (20) phrase addition. Features of these types were coded by location within each record, and a sample listening assessment document for one speaker is attached in an appendix. A comparable text-file document for each speaker will be distributed with the full database. Any records or portions of records passed without comment should be considered unremarkable. It should be no surprise, in a dataset as large as this, that not all records contain exactly what one might assume from the content of the printed task list used to

elicit performance.   That is, only a subset of records are ideally "clean."  It is also important to remember that microbeam experiments are done "live," with little rehearsal.  Except in unusual circumstances, re-takes of Database records poorly performed were not routinely attempted. Concern about performance precision is secondary to concern for the physical risk to subjects inherent in the microbeam technique.  However, the presence of a variety of "normal" speech errors should not be viewed as a liability, given our goal of sampling some approximation of the varieties of natural speech behavior, and our wish to sample speech behaviors broadly (see section 3.1 above).  The natural inclusion of some errors provides analysis opportunities (and views of variability) unavailable in more precisely controlled data.

# CHAPTER FOUR
## Speaker Sample

Speaker samples greater than three of four are unusual in speech physiological research. This is probably true for practical reasons. Reducing multiple signal streams in the absence of well-founded methods takes a great deal of time, and returns on that investment are wildly uncertain. Sometimes, too, methods of acquiring data are risky and extraordinarily difficult.

Production studies with more than just acoustic data, and "large" speaker samples (N > 1!), are especially problematic from a "methods" point of view. This is because they force us to face the problem of whether and how to average across individual performances. And even when generalizations are successfully drawn, we are left with the sticky problem of explaining why individual differences exist. Unfortunately, the possibilities are many. Physically, speakers are big or little, weak or strong, red-headed or blonde. Mentally, they are sane or mad, cranky or calm. They grow up in different places, eat different foods, shave often or not enough, and -- some say -- exercise free will. It is human nature to want to know why people faced with the same task behave differently. It is the behaviorist's curse to try to work it out.

Toward this end, we have assembled a modest amount of supplemental information on each of our speakers. In several ways, this information may afford insight regarding individual differences in speech kinematic behavior. The idea that speakers move as they do partly because of the ways they are built is not uncommon. This natural inference seems directly supported, for example, by data published by Kuehn & Moll (1976), showing strong positive relationships between tongue and jaw displacement and velocity, and tongue or jaw size, for a sample of five speakers pre-selected to span a wide range of structural sizes. A more recent report of Edwards & Harris (1990) is conceptually similar to the earlier work of Kuehn & Moll, in suggesting that degree of independence of mandibular rotation and translation during speech production may be linked to a speaker's occlusal classification.

In the paragraphs, tables, and figures that follow, we summarize various types of supplementary information for database speakers.

### 4.1 Recruitment and screening

Speakers represented in the database were recruited by advertisement from the campus of the University of Wisconsin-Madison and from the surrounding city. Potential speakers were screened for general health, speech-language characteristics, hearing, oral-motor structure and function, oral reading ability, and willingness to endure the rigors of the experimental protocol. All were young adults, with no evident neuromotor or articulation disorders, and all were in good general health, as judged from direct observation and a standard set of questions regarding present and past health. None possessed a self-reported history of speech-language pathology, nor current evidence of pathology as judged by a certified speech-

language pathologist. All passed a puretone hearing screening with thresholds at or below 25dB HL (1989 ANSI standard), at frequencies of 500 Hz, 1 kHz, 2 kHz, 3 kHz, 4 kHz, and 8 kHz. A conventional oral peripheral mechanism exam was used to judge oral-motor structure and function. During data collection, speech was elicited through orthographic presentation of tasks on a CRT. Therefore, speakers were expected to demonstrate a minimally acceptable degree of fluent oral reading ability, at the paragraph level, as informally judged by the examiner. This was to ensure that difficulties with oral reading (of whatever sort) would not interfere with the Database Project goal of collecting normal connected speech samples which approximated (as nearly as practicable under the circumstances) speech kinematic behaviors as they might exist in nonexperimental conditions.

Potential speakers were informed at length of the nature of the project, data collection procedures, and potential risks associated with their participation. They were actively and repeatedly discouraged from participation if they had any reservations whatever. All potentially willing speakers were required to take a minimum of 24 hours to consider their participation before consenting. Speakers were paid $5 for attendance at the initial screening, and $10/hr for their subsequent participation in data collection. All procedures were approved by an Institutional Review Board of the University of Wisconsin.

*4.2 Demographics*

Usable data exist for 57 different speakers, 32 females and 25 males. Median age for the speaker sample was 21.1 years, with quartiles at 20.1 and 23.9 years. Median ages among females and males were 21.3 and 20.8 years, respectively, with quartiles at (20.16, 24.7) and (20.0, 22.4), and ranges of (18.3, 37.0) and (18.6, 29.3) years, also respectively. Speaker age from date of birth to run date, is included in Table 4.1, along with supplemental information including height, weight, educational level, and non-English language training. The restricted age range of the speaker sample is easy to understand given that our primary source population was the UW student body. An additional constraint of the microbeam technique, that speakers be largely free of metal dental fillings, also biased our sample toward relatively young participants.

*4.3 Dialect and residency*

A majority of speakers spoke an Upper Midwest dialect of American English. Thirty (30), in fact, may reasonably be termed native sons or daughters of Wisconsin. Speaker dialect was defined by responses to the questions "Where did you grow up?" and/or "Where did you live between the ages of roughly 4-12 years?". *Dialect Base* (i.e., place of residence during linguistically formative years) was distinguished both from *Place of Birth* and *Permanent Residence* (i.e., at the time of data collection). Place of birth, dialect base, and permanent residence are coded by both city and state in Table 4.2. Figure 4.1 graphically presents the geographical distribution of speakers' dialect base. Questions about permanent residence, for students at most public universities, often elicit a predictable response. For tuition purposes, state-supported universities insist that students still live permanently wherever their parents live. Students, on the other hand, typically insist that this is nonsense, and that they live permanently

somewhere their parents do **not** live, usually in the university community itself. In this regard, our database records are as much a victim of residency politics as are voter rolls of many states.

| Dataset | Gender | Age (yrs) | Height (in) | Weight (lbs) | Education | Languages |
|---|---|---|---|---|---|---|
| JW05 | F | 22.41 | 65 | 120 | HS + 2 yrs. | Spanish(HS/4) |
| JW06 | M | 24.59 | 70 | 155 | BA + 1 yr. | Spanish(C/2) |
| JW07 | M | 27.5 | 73 | 185 | BA | French, Russian, German(1 yr. of each) |
| JW08 | M | 22.83 | 70 | 160 | BA | French(HS); Russian(C) |
| JW09 | F | 30.36 | 67 | 200 | HS + 7 yrs. (pt tm) | French(HS) |
| JW11 | M | 20.04 | 69 | 150 | BS | Spanish(HS); French(HS) |
| JW12 | M | 21.1 | 72 | 155 | HS + 3 yrs. | Spanish(HS) |
| JW13 | F | 36.98 | 65 | 165 | HS | French(HS) |
| JW14 | F | 36.02 | 68 | 145 | BS | Spanish(HS) |
| JW15 | M | 22.44 | 66 | 130 | HS + 4 yrs. | Spanish(HS/4yrs) |
| JW16 | F | 20.49 | 65 | 140 | HS + 2 yrs. | German(HS/3; C/0.5); some German spoken in home |
| JW17/22 | F | 20.65 | 66 | 125 | HS + 2 yrs. | French(HS/3) |
| JW18 | M | 19.38 | 69 | 170 | HS + .5 yr. | Spanish(HS/4) |
| JW19 | M | 21.66 | 70 | 155 | HS + 3 yrs. | Spanish(HS/4) |
| JW20 | F | 25.39 | 67 | 125 | BS + 1 yr. | French(HS) |
| JW21 | F | 21.57 | 66 | 140 | HS + 3 yrs. | Spanish(HS/3); French(C/1) |
| JW23 | F | 25.12 | 64 | 117 | HS + 3 yrs. | Cantonese(semifluent);Russian(4 yrs Col);Turk(1);Japanese(1) |
| JW24 | M | 19.65 | 72 | 150 | HS + 1 yr. | Russian(C/1.5) |
| JW25 | F | 24.19 | 65 | 125 | BS | none |
| JW26 | F | 23.88 | 65 | 140 | BA | French(fluent: HS+C+1 yr abroad); Spanish(C/1) |
| JW27 | F | 20.85 | 65 | 190 | HS + 2 yrs. | French(HS+C/4 total) |
| JW28 | M | 21.58 | 73 | 165 | HS + 2 yrs. | German(HS/4;C/1;exch. student x2) |
| JW29 | F | 20.61 | 66 | 125 | HS + 2 yrs. | Spanish(HS/3) |
| JW30 | F | 19.45 | 65 | 150 | HS + 1 yr. | German(HS+Mid/5; 1 year in Germany @ age 7) |
| JW31 | F | 19.92 | 62 | 130 | HS + 1 yr. | German(HS/2) |
| JW32 | M | 21.85 | 72 | 175 | HS + 2 yrs. | Spanish(HS/2) |
| JW33 | F | 19.41 | 66 | 125 | HS + 1 yr. | French(C?/2); Chinese(HS/3) |
| JW34 | F | 20.99 | 68 | 145 | HS + 2 yrs. | French(HS/1;C/2) |
| JW35 | F | 24.28 | 71 | 165 | BA | Spanish(HS/3); French(C/1.5) |
| JW37 | F | 20.12 | 67 | 120 | HS + 2 yrs. | Spanish(HS/4) |
| JW38/2 | F | 20.88 | 65 | 120 | HS + 3 yrs. | Spanish(HS/5); Italian(C/1.5) |
| JW36 | F | 18.33 | 62 | 110 | HS | Spanish(HS/4;C/0.5) |
| JW39 | F | 23.77 | 66 | 140 | BA | Spanish(HS/5;C/2) |
| JW40 | M | 19.21 | 73 | 165 | HS + 1 yr. | German(HS/2); Japanese(C/0.5) |
| JW41 | M | 20.41 | 75 | 190 | HS + 2 yrs. | Spanish(HS/2); German(C/2) |
| JW42 | M | 18.56 | 73 | 165 | HS | Spanish(HS/3.5) |
| JW43 | M | 20.05 | 74 | 180 | HS + 1 yr. | French(HS/5;C/0.5) |
| JW44 | M | 20.75 | 69 | 170 | HS | French(HS/5) |
| JW45 | M | 21.24 | 69 | 173 | HS + 3 yrs. | German(HS/5) |
| JW46 | F | 19.57 | 63 | 110 | HS + 2 yrs. | Spanish(HS/4;C/1.5) |
| JW47/2 | F | 25.32 | 71 | 160 | BS | French(C/1) |
| JW48 | F | 21.33 | 61 | 150 | HS | None |
| JW49 | F | 19.45 | 63 | 115 | HS + 1 yr. | Spanish(HS/4;C/1) |
| JW502 | F | 34.04 | 62 | 175 | BA/BS | Spanish (HS/2) |
| JW51 | M | 19.19 | 72 | 175 | HS | Spanish(HS) |
| JW52 | F | 26.46 | 68 | 125 | BA | French(HS+C/6 total) |
| JW53 | M | 20.72 | 72 | 155 | HS + 2 yrs. | German(HS+C/7 total); Chinese(C/1) |
| JW54 | F | 21.34 | 67 | 120 | HS + 2 yrs. | None |
| JW55 | M | 26.71 | 72 | 182 | BA | Spanish(HS;C); French(C/0.5) |
| JW56 | F | 22.33 | 67 | 140 | BS | Spanish(HS/3;C/0.5); 1 summer in Brazil(Port.); Czech in home |
| JW57 | M | 19.6 | 74 | 195 | HS + 1 yr. | Spanish(HS/3); French(C/1.5) |
| JW58 | M | 23.23 | 71 | 165 | BA | Hebrew(6+;C/1.5) |
| JW59 | M | 29.28 | 69 | 160 | MA | German(HS/4;C/2) |
| JW60 | F | 20.19 | 67 | 130 | HS + 2 yrs. | French(HS/4) |
| JW61 | M | 20.4 | 70 | 150 | HS + 2 yrs. | French(HS/2;C/2) |
| JW62 | F | 18.38 | 64 | 110 | HS | Spanish(HS/2) |
| JW63 | M | 20.66 | 71 | 145 | HS + 2 yrs. | French(HS/3); Italian(C/1) |

Table 4.1: Demographics

| Table 4.2: Place of Birth (POB), Dialect Base (DB), and Residence (Res) | | | | | |
| --- | --- | --- | --- | --- | --- |
| Dataset | POB City | POB State | DB City | DB State | Res City | Res State |
| JW05 | St. Louis | MO | St. Louis | MO | St. Louis | MO |
| JW06 | Akron | OH | Chicago | IL | Madison | WI |
| JW07 | Maquoketa | IA | Maquoketa | IA | Monona | WI |
| JW08 | Milwaukee | WI | Brookfield | WI | Madison | WI |
| JW09 | Ames | IA | Whitewater | WI | Madison | WI |
| JW11 | St. Louis | MO | Brookfield | WI | Matthews | NC |
| JW12 | Marinette | WI | Marinette | WI | Appleton | WI |
| JW13 | Rockford | IL | Rockford | IL | Madison | WI |
| JW014 | Neilsville | WI | Stoughton | WI | Stoughton | WI |
| JW15 | Milwaukee | WI | Milwaukee | WI | Madison | WI |
| JW16 | Chilton | WI | Kiel | WI | Madison | WI |
| JW17/22 | Mankato | MN | Mankato | MN | Rochester | MN |
| JW18 | Hudson | WI | Hudson | WI | Hudson | WI |
| JW19 | Detroit | MI | Detroit | MI | Boca Raton | FL |
| JW20 | Milford | MA | Milford | MA | Madison | WI |
| JW21 | Philadelphia | PA | Cherry Hill | NJ | Cherry Hill | NJ |
| JW23 | Sacramento | CA | Hong Kong | | Madison | WI |
| JW24 | Ft. Atkinson | WI | Jefferson | WI | Madison | WI |
| JW25 | Hillsboro | WI | Elroy | WI | Madison | WI |
| JW26 | Mt. Kisco | NY | Verona | WI | Madison | WI |
| JW27 | Blair | WI | Blair | WI | Blair | WI |
| JW28 | Madison | WI | Madison | WI | Rhinelander | WI |
| JW29 | Milwaukee | WI | Milwaukee | WI | Milwaukee | WI |
| JW30 | Edina | MN | Edina | MN | Edina | MN |
| JW31 | Chilton | WI | New Holstein | WI | New Holstein | WI |
| JW32 | Souix City | IA | Huron | SD | La Crosse | WI |
| JW33 | Buena Vista | CO | Minneapolis | MN | Minneapolis | MN |
| JW34 | Amery | WI | Amery | WI | Amery | WI |
| JW35 | Beaver Dam | WI | Waupon | WI | Waupon | WI |
| JW36 | Blue Island | IL | Park Forrest | IL | Park Forrest | IL |
| JW37 | Los Angeles | CA | Morgan | CA | Ithaca | NY |
| JW38/2 | Brooklyn | NY | Great Neck | NY | Great Neck | NY |
| JW39 | Houston | TX | Rochester | MN | Madison | WI |
| JW40 | Appleton | WI | Green Bay | WI | Green Bay | WI |
| JW41 | Milwaukee | WI | Milwaukee | WI | Milwaukee | WI |
| JW42 | Milwaukee | WI | Greendale | WI | Greendale | WI |
| JW43 | Wauwatosa | WI | Waukesha | WI | Waukesha | WI |
| JW44 | Madison | WI | Madison | WI | Madison | WI |
| JW45 | La Porte | IN | Mishawaka | IN | Madison | WI |
| JW46 | Skokie | IL | Buffalo Grove | IL | Buffalo Grove | IL |
| JW47/2 | Des Moine | IA | Indianola | IA | Madison | WI |
| JW48 | Maywood | IL | Maywood | IL | Madison | WI |
| JW49 | Ashland | WI | Madison | WI | Madison | WI |
| JW502 | Waukesha | WI | Madison | WI | Madison | WI |
| JW51 | Madison | WI | Madison | WI | Madison | WI |
| JW52 | Scarborough | Ontario | Kewanee | IL | Madison | WI |
| JW53 | Ft. Worth | TX | Waukesha | WI | Wauwatosa | WI |
| JW54 | Wonewoc | WI | Wonewoc | WI | Madison | WI |
| JW55 | Witchita | KS | Denver | CO | Madison | WI |
| JW56 | Minneapolis | MN | Edina | MN | Edina | MN |
| JW57 | Racine | WI | Union Grove | WI | Union Grove | WI |
| JW58 | Edison | NJ | Fair Lawn | NJ | Chicago | IL |
| JW59 | Sauk City | WI | Sauk city | WI | Madison | WI |
| JW60 | Fairfax | VA | Columbus | OH | Westchester | OH |
| JW61 | Ann Arbor | MI | Middleton | WI | Middleton | WI |
| JW62 | Washington, DC | | Crofton | MD | Verona | WI |
| JW63 | Philadelphia | PA | Los Angelos | CA | Los Angeles | CA |

**Figure 4.1:** Dialect base distribution
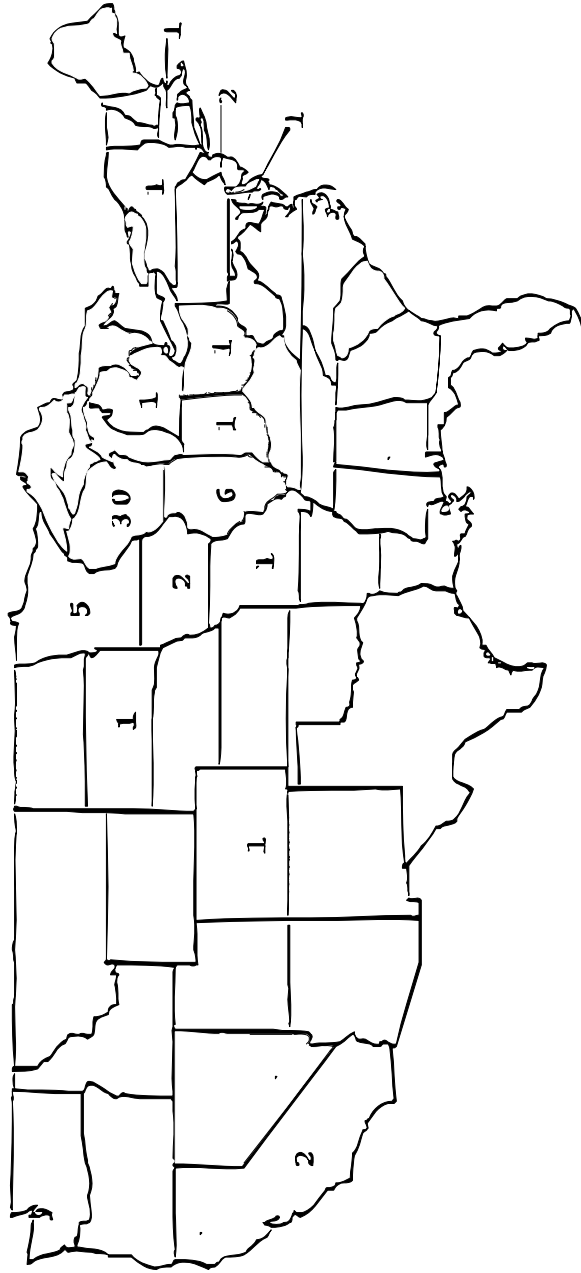


Hong Kong: 1

*4.4 Dentition*

From each speaker we obtained plaster models fabricated from impressions of the upper and lower dental arches. These models served numerous purposes. In addition to providing sources of dental information summarized here, they were used to estimate palatal depth, to derive palatal outlines (see section 5.2.2.4), and to fabricate speaker-specific calibration biteplates (see section 5.2.2.1). Measures of dental arches are potentially useful because they delineate the size of the cavity within which the tongue must work during speech.

*4.4.1 Occlusion and missing teeth*

Table 4.3 describes occlusion and missing teeth for each speaker. Occlusal class by Angle's rule (Angle, 1907) was determined from articulated maxillary and mandibular arches of the plaster models. Missing teeth are noted using the Universal Numbering System (for example, see Short, 1987).

*4.4.2 Dental Measures*

Figure 4.2 (right half) summarily illustrates maxillary arch measures. These include the width and length of the maxillary arch at three dental landmarks: the canines, the mesial-buccal cusp tip of the first molar, and the distal-buccal cusp tip of the second molar. Articulated models were also used to determine extent of incisal overjet and overbite (these values are stored in ASCII files with other Database materials). Palatal depths measured from the maxillary models, and from two-dimensional palatal outlines based on these models, are also summarily illustrated in Figure 4.2 (left side). The figure shows the range of depth measures at each of the three dental landmarks described above. It also shows three contrasting exemplars of palatal shape, to suggest some of the structural variability in our subjects. As with all midsagittal plane representations we provide, the origin of the coordinate system is at the tip of the maxillary incisors, and the x-axis is defined as the maxillary occlusal plane (see section 5.2.2.1).

*4.4.3 Tempero-mandibular joint (TMJ) condition and dental health*

Table 4.4 summarizes coarsely observable characteristics of TMJ behavior, and subject self-reports about such behavior. The same table summarizes salient information about oral surgery and orthodontia, also based on subject reports.

*4.5 Anthropomorphic Measures*

Measures of cranial and mandibular size and shape were recorded for each speaker, following standard anthropomorphic methods described by Farkas (1981). Figure 4.3 graphically illustrates distributions of these measures. Mandibular measures were derived from approximations of projected mid-sagittal-plane coordinates of the superior edge of the central mandibular incisors, gnathion, gonion, condyle center, and coronoid process, relative to cranial reference axes. Figure 4.4 summarizes these results graphically. All numerical values are also stored in ASCII files that accompany other database materials.

| Table 4.3: Dental information I | | | | | |
|---|---|---|---|---|---|
| Dataset | Occlusion | Missing Molars | Missing Pre-molars | Missing Canines | Missing Incisors |
| JW05 | normal | 1,16,17,32 | 5,12,21,28 | none | none |
| JW06 | normal | 1,16,17,32 | 21 | none | none |
| JW07 | normal | 1,16,17,32 | none | none | none |
| JW08 | normal | 1,16,17,32 | none | none | none |
| JW09 | normal, crossbite on right | 1,16,17,32 | none | none | none |
| JW11 | normal, bilateral post. crossbite | 1,16,17,32 | none | none | none |
| JW12 | normal left, class II right, openbite | 1,16,17,32 | 5 | none | none |
| JW13 | class II | 1,16,17,32 | none | 6,11 | none |
| JW14 | normal | 1,16,17,32 | 5,12,21,28 | none | none |
| JW15 | normal | 1,16,17,32 | none | none | none |
| JW16 | normal | 1,16,17,32 | none | none | none |
| JW17/22 | class II | 1,16,17,32 | 5,12 | none | none |
| JW18 | normal | 1,16,17,32 | none | none | none |
| JW19 | class III | 1,16,17,32 | 20,29 | none | none |
| JW20 | normal, openbite, incisors edge-to-edge | 1,16,17,32 | none | 6 | none |
| JW21 | normal | 1,32 | 5,12,21,28 | none | none |
| JW23 | normal | 1,16,17,32 | none | none | none |
| JW24 | normal | 1,16,17,32 | none | none | none |
| JW25 | normal | 1,16,17,32 | none | none | none |
| JW26 | normal | 1,16,17,32 | none | none | none |
| JW27 | normal | 1,16,17,32 | none | none | none |
| JW28 | normal, unilat. crossbite on left | 1,16,32 | none | none | none |
| JW29 | normal | 1,16,17,32 | none | none | none |
| JW30 | normal | none | none | none | none |
| JW31 | normal | 1,16,17,32 | none | none | none |
| JW32 | class II, crossbite | none | none | none | none |
| JW33 | normal | 1,16,17,32 | none | none | none |
| JW34 | normal | 1,16,17,32 | none | none | none |
| JW35 | class II, crossbite | 1,16,17,32 | none | none | none |
| JW36 | normal | 1,16,17,32 | none | none | none |
| JW37 | normal, unilat. post. openbite | 1,16,17,32 | none | none | none |
| JW38/2 | normal | 1,16,17,32 | none | none | none |
| JW39 | normal | 1,16,17,32 | none | none | none |
| JW40 | normal, crossbite on right | 1,16,17,32 | none | none | none |
| JW41 | normal | 1,16,17,32 | 5,12,21,28 | none | none |
| JW42 | normal | 1,16,17,32 | none | none | none |
| JW43 | normal, ant. openbite | 1,16,32 | none | none | none |
| JW44 | normal | 1,16,17,32 | none | none | none |
| JW45 | normal | 1,16,17,32 | none | none | none |
| JW46 | normal | 1,16,17,32 | none | none | none |
| JW47/2 | normal | 1,16,17,32 | none | none | none |
| JW48 | Class II | none | none | none | none |
| JW49 | normal | 1,16,17,32 | none | none | none |
| JW502 | normal, bilat. crossbite | 1,16,17,32 | none | none | none |
| JW51 | normal right, class II left, crossbite | none | none | none | none |
| JW52 | normal right, class II left | 1,16,17,32 | none | none | none |
| JW53 | normal | 1,16,17,32 | none | none | none |
| JW54 | normal | 1,16,17,32 | none | none | none |
| JW55 | normal | 1,16,17,32 | none | none | none |
| JW56 | normal | 1,16,17,32 | none | none | 23 |
| JW57 | normal, unilat. right crossbite | 1,16,17,32 | none | none | none |
| JW58 | normal | 1,16,17,32 | none | none | none |
| JW59 | normal | 1,16,17,32 | none | none | none |
| JW60 | normal | 1,16,17,32 | none | none | none |
| JW61 | Class II | 1,16,17,32 | 5,12, | none | none |
| JW62 | normal | 1,16,17,32 | 5,12,21,28 | none | none |
| JW63 | normal right, crossbite left | 1,16,17,32 | none | none | none |

Table 4.4: Dental information II

| Dataset | TMJ status | Oral surgery | Orthodontia |
|---|---|---|---|
| JW05 | bilateral clicking with wide opening; easy subluxation | 3rd molars ext @ 16 yrs | @ 11-13 yrs 1st premolars ext |
| JW06 | NA | 3rd molars ext @ unspecified date | @ 15-17 yrs |
| JW07 | NA | 3rd molars ext @ 16 yrs | @ 13-15 yrs |
| JW08 | bilateral popping | 3rd molars ext @ 22 yrs | @ 12-14 yrs |
| JW09 | Bilateral clicking,episodic superficial pain | 3rd molar @ 20 yrs | @ 13-15 yrs |
| JW11 | NA | NA | NA |
| JW12 | reported assymetric click not apparent; easy subluxation | Unspecified dental work @ 7 yrs | NA |
| JW13 | NA | 3rd molars ext; unspecified date | @ 16-18 yrs, upper lat incisor and canine ext |
| JW14 | NA | 3rd molars ext @ 24 yrs | @ 16 yrs, 1st premolars ext |
| JW15 | NA | NA | NA |
| JW16 | NA | 3rd molars ext @ 17yrs | NA |
| JW17/22 | NA | 3rd molars ext @ 17 yrs | @ 13-15 yrs; 1st premolars ext |
| JW18 | bilateral clicking with opening of ca. 1 cm | NA | retainer in middle school, 2 yrs |
| JW19 | NA | lower premolar ext @ 5 yrs | @ 12-13 yrs to correct overbite |
| JW20 | Bilateral clicking, intermittent pain w/ stress | Rt max canine ext @ 12 yrs; 3rd molars @ 21 yrs | @ 7-8 & 11-12 yrs |
| JW21 | easy sublaxation, otherwise assymptomatic | 3rd molar, rt side, ext @ 15 yrs | @ 11-14 yrs. 1st premolars ext |
| JW23 | NA | 3rd molars ext @ 20 yrs | NA |
| JW24 | NA | 3rd molars ext @ 18 yrs | NA |
| JW25 | NA | NA | NA |
| JW26 | NA | 3rd molars ext @ 18 yrs | NA |
| JW27 | irregular clicking | NA | NA |
| JW28 | assymetric clickling w/ wide opening | NA | NA |
| JW29 | NA | NA | @ 13-14 yrs |
| JW30 | NA | NA | @ 14-15 yrs |
| JW31 | intermittent clicking in mastication | NA | @ 13-16 yrs; ext man 3rd molars, max 2nd molars |
| JW32 | NA | NA | NA |
| JW33 | NA | NA | NA |
| JW34 | NA | NA | retainer @ 14-15 yrs |
| JW35 | assymetric opening; unilateral clicking (lt) | NA | NA |
| JW36 | easy subluxation w/ wide opening | NA | NA |
| JW37 | NA | 3rd molars ext @ 15 yrs | @ 12-14 yrs permanent retainers |
| JW38/2 | NA | NA | NA |
| JW39 | NA | NA | NA |
| JW40 | NA | 3rd molars ext @ 18 yrs | @ 13-15 yrs |
| JW41 | NA | 3rd molars ext @ 19 yrs | @ 14-16 yrs 1st premolars ext |
| JW42 | NA | NA | @ 11-14 yrs |
| JW43 | assymetric (lt) clicking, even with small movement | NA | NA |
| JW44 | assymetric clicking (rt) w/ moderate opening | NA | NA |
| JW45 | NA | 3rd molars ext @ 17 yrs | NA |
| JW46 | NA | 3rd molars ext @ 18 yrs | NA |
| JW47/2 | NA | Rt max canine ext @ 10 yrs | NA |
| JW48 | NA | NA | NA; poor hygiene |
| JW49 | NA | NA | NA |
| JW502 | minor rt side clicking @ narrow opening, 1cm | 3rd molars ext @ 17 yrs | NA |
| JW51 | NA | NA | NA |
| JW52 | NA | NA | NA |
| JW53 | bilateral clicking with wide opening; no pain | tumor removal @ 6 yrs | headgear @ 9-10 yrs |
| JW54 | NA | 3rd molars ext @ 20 yrs | NA |
| JW55 | NA | 3rd molars ext @ 26 yrs | NA |
| JW56 | NA | 3rd molars ext @ 20 yrs | @ 13-15 yrs; ext man lat incisor |
| JW57 | NA | 3rd molars ext @ 15 yrs | NA |
| JW58 | NA | 3rd molars ext @ 17 yrs | @ 10-12 yrs |
| JW59 | NA | 3rd molars ext @ 27 yrs | NA |
| JW60 | bilateral clicking with wide opening | NA | @ 17-18 yrs |
| JW61 | NA | 3rd molars ext @ 19 yrs | @ 14-17 yrs; max premolars ext |
| JW62 | NA | 3rd molars ext @ 17 yrs | @ 12-14 yrs; premolars ext |
| JW63 | NA | NA | @ 13-15 yrs |

Figure 4.2: Palate and maxillary arch measures



Maxillary Arch Dimensions
(mm)

34.8

50.9

59.2

Palate Height
(mm)

C

M1

M2

Position wrt MaxOP (mm)
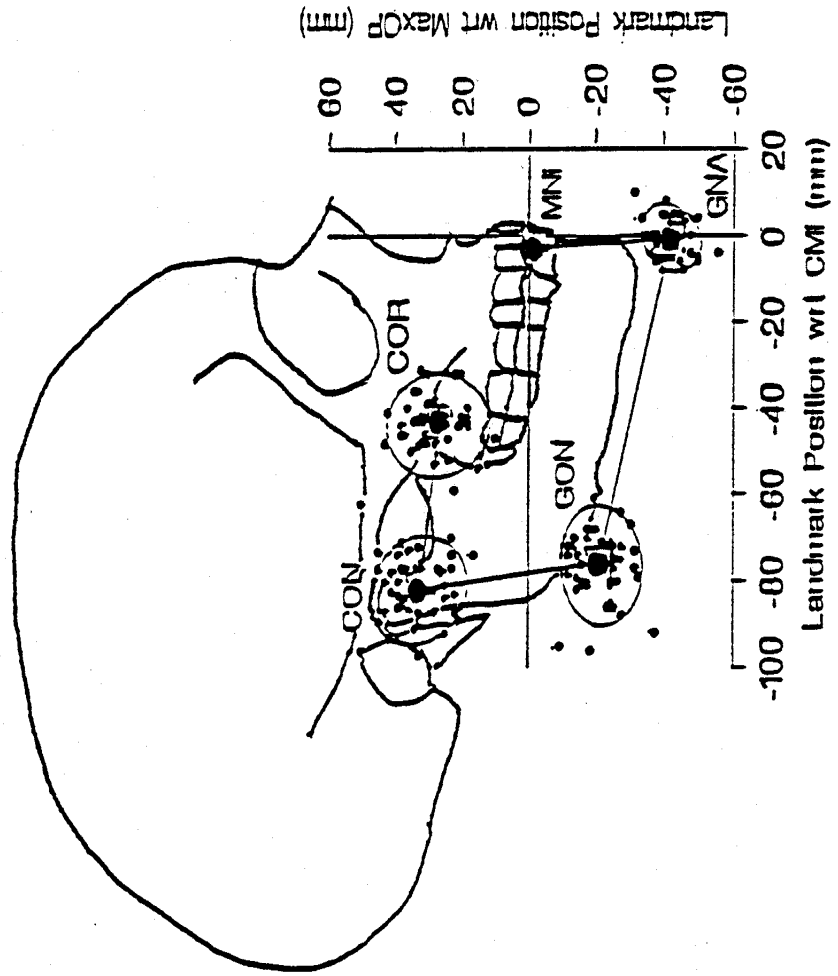
Position wrt CMI (mm)
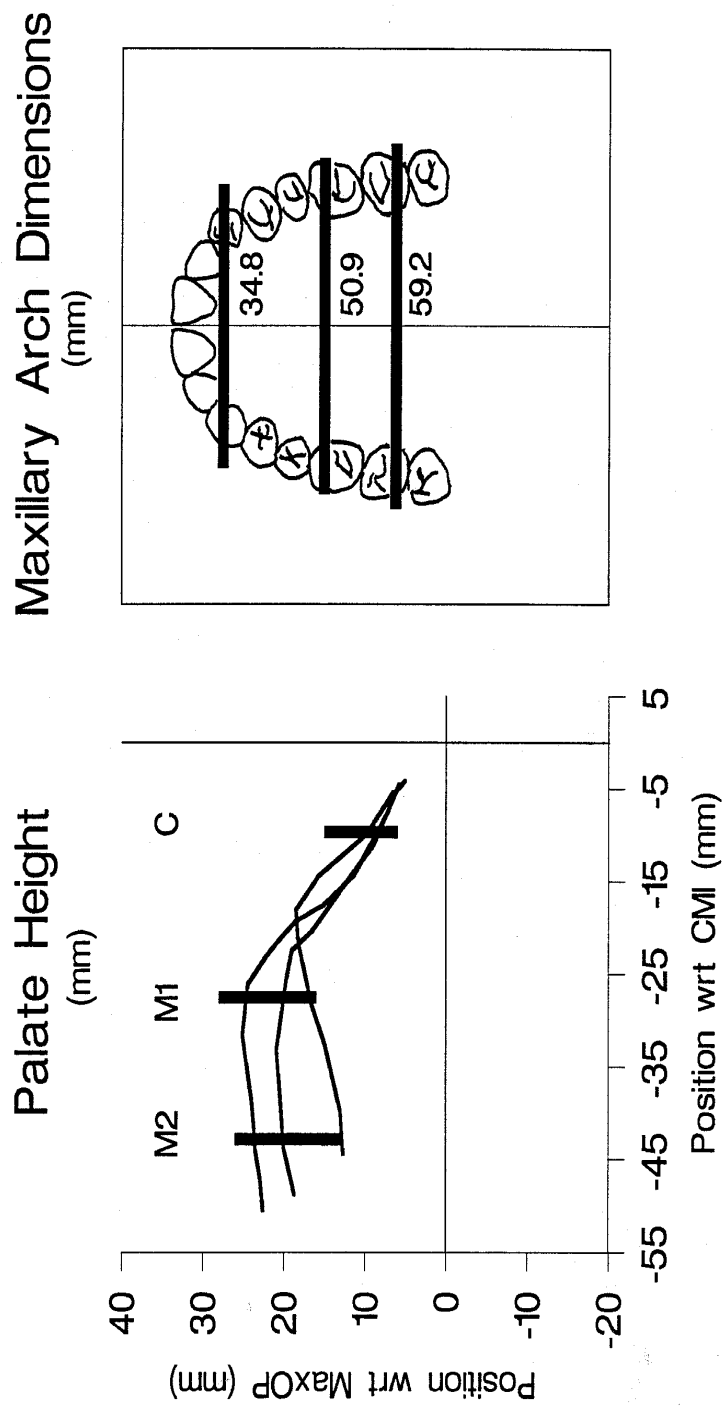
# Figure 4.3: Cranial measures
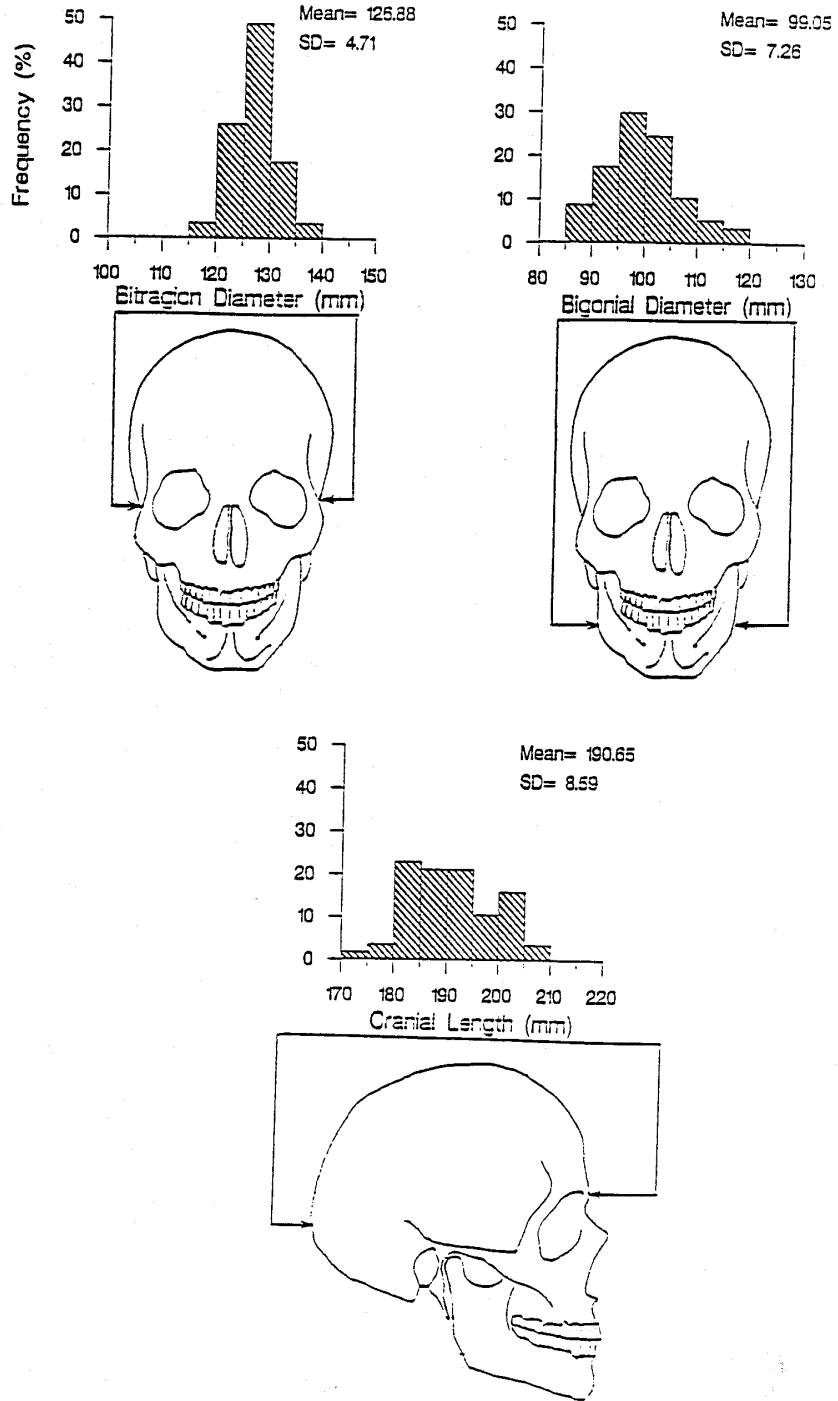
Figure 4.4: Distribution of mandibular landmarks

# CHAPTER FIVE
## Experimental Protocol

Developing a database is a complicated task, and it is difficult to guess the full significance of all procedures for data obtained. In sections that follow, we have summarized the salient features of our experimental protocol, in the hope that this information will enhance the value of the resulting signal streams and any relevant supplementary materials.

### 5.1. Data Types

A common data acquisition protocol was followed for each speaker, in which three types of time-series data were recorded: (1) wide-band physiological tracks; (2) videophotographic images; and, (3) low-band pellet position tracks. The purposes, methods, and sampling and storage conventions for each data type are described separately.

### 5.1.1. Wide-band physiological data

By convention, two channels recorded during each experimental session are referred to as wide-band *physiological* data: the radiated sound pressure wave, and a representation of neck wall vibration overlying the thyroid lamina. The XRMB facility was developed to enhance understanding of the relationship between articulatory movements, the resulting speech wave, and other biological events associated with speech. Thus, signal streams of the latter types are recorded concurrently with representations of articulatory movements.

### 5.1.1.1. Sound Pressure Wave (SPW)

The SPW was sensed by a directional microphone (Shure SM81 Condenser) positioned at mouth level, roughly 10 cm anterior and 5 cm lateral to the mouth opening. The microphone signal was fed in parallel to three separate devices. The first of these was a 15-bit-resolution A/D converter programmed to digitally sample the SPW at 21379 times per second, and to store the resulting digital stream synchronously with pellet position histories, on SMD computer disks, in the XRMB archive. Prior to digital conversion, an anti-aliasing filter (-3dB at 7500 Hz) was applied to the microphone signal. The digital conversion rate of 21739 samples/second applied to the SPW was selected because it provides sufficient bandwidth for most speech-acoustic analyses; is closely compatible with the output rate of an inexpensive, commercially-available D/A board for personal computers (IBM: M-ACPA); and, is as close as practically possible to a convenient rational proportion (ca. one-half) of the 44.1 kHz commercial/industrial standard for DAT recorders. For two speakers only, early in the sample, JW7 and JW8, the AD conversion rate for the SPW was 16129 samples/second, approximately three-fourths of the 21739-rate used for all other speakers, and roughly three-eighths of the 44.1 kHz DAT standard.

The second device receiving the microphone output was a studio-quality two-channel DAT recorder [Sony PCM-2500], which digitally sampled the SPW at 44100 times per second (without the recorder's pre-emphasis function, or any pre-filtering other than that intrinsic to the

recorder's own circuitry).  The resulting signal was digitally stored on DAT-type magnetic mastering tape.

The third device receiving the microphone output was a video recorder (see Section 5.1.2), which captured the audio signal synchronously with video images, as a "banter" channel, on standard magnetic videocassette tape.

### 5.1.1.2. Neck Wall Vibration (NWV)

A piezoelectric surface-mount accelerometer with built-in FET preamplifier (Knowles Electronics, model BU-1771; frequency response:  50-3000 Hz), snugly taped to the skin surface overlying the thyroid lamina, was used to obtain a representation of neck wall vibration associated with phonation.  The accelerometer signal was recorded in anticipation of future interests in pitch-synchronous LPC analysis of the speech acoustic wave.  Its output was fed in parallel to both the A/D converter, and DAT recorder.  For most speakers, the A/D converter was programmed to digitally sample the NWV signal at 5434 times per second, one-quarter as frequently as the SPW, and to store the resulting digital stream synchronously with the SPW and pellet position histories, on computer disk.  For two speakers, JW7 and JW8, the NWV was sampled at 5376 times per second, one-third as frequently as their respective SPW sampling rates; for one speaker, JW9, the NWV was sampled 4831 times per second, two-ninths as frequently as her SPW.  For all speakers, the DAT recorder digitally sampled the NWV at 44100 times per second, and stored the resulting signal stream on a second channel on magnetic mastering tape, synchronously with the DAT representation of the SPW.

### 5.1.2. Videophotographic images

Synchronous lateral and frontal video images were recorded at 60 fields/second, split-field fashion, onto standard video tape, using a Sanyo VHR 8310 recorder with a Vidicraft SEG-200. Video images were obtained on line to monitor speakers' positions in the microbeam image field. They were also recorded to provide a banter record of each experiment, and have proven useful in understanding certain speaker movements that affect data accuracy.  The image quality is not high, however, and it is unlikely that the images would be of future use to investigators interested, for example, in speech reading.

### 5.1.3. Low-band pellet position histories

The XRMB system was used to digitally track the motions of 2.5-3.0 mm gold pellets attached to the head (3 pellets, as reference or fiducial markers), upper and lower lips (1 pellet each), tongue surface (4 pellets), and mandible (2 pellets), during each speaker's performance of records from the task inventory.  The pellet position histories, reflecting discrete fleshpoints and bony landmarks on articulators, form our representation of the pattern of movements associated with each task.  A brief summary of the XRMB method for tracking pellets is reproduced from Westbury (1991):

> *The system performs its primary, pellet-tracking function... following a general method invented and first implemented by Fujimura and his colleagues at the University of Tokyo (cf., Fujimura, Kiritani, & Ishida, 1973; Kiritani, Itoh, & Fujimura, 1975).  An electron*

31

*beam accelerated by a voltage source of up to 600 kV, at a current of up to 5 mA, is focused on a tungsten target to produce X-rays. A narrow beam of the incident X-rays, roughly 0.4 mm in diameter, then passes through a pinhole aperture, with the path of the X-ray beam determined by the location of the electron beam on the target. A series of high-speed computations, partly within dedicated digital circuits, is used to adjust the X-ray beam path from moment to moment, to produce local X-ray scans roughly 6 mm square, circumscribing the expected location of each pellet within the system image field. The position of a pellet is determined when it falls within such a scan area, and produces a recognizable "shadow" on a two-dimensional X-ray count registered by an NaI (sodium-iodide) crystal detector. Current and previous positions of a pellet are used to predict both its future position along its trajectory, and the required location of a subsequent local scan of the X-ray beam. The cycle of operations, including local scan, recognition, and prediction, is repeated for each of the pellets, and the entire cycle of all operations is repeated for as many as [twelve] pellets at an aggregate rate of [approximately] 700 times per second. The frequency of repetition for each pellet is specified separately, at rates ranging between [20 and] 180 times per second.*

*At discrete moments in time, the system assigns rectangular coordinates to the evaluated centers of pellets, by approximating the centroids of their respective shadows. These coordinates for any specific pellet and moment in time represent the image-plane projection of the pellet's position, and are proportional to the known electron beam deflections that are required to generate the X-ray beam whose path the pellet interrupts.*

Time-stamped coordinates for each pellet, for each of its sampled positions, were stored on computer disk, in synchrony with SPW and NWV digital signal streams. For a variety of reasons, the aggregate pellet-sampling rate for real-time tracking is limited to approximately 700 samples per second. This constraint relates primarily to the amount of time the electron beam is on any specific pixel of the tungsten target. A dwell time of 10 microseconds was typically used, representing a compromise dictated by the target's heat-duty-cycle, the need for sufficient per-unit-area photon density, and the customary sizes of the tracking rasters that were generated to follow pellets (10x10 pixels for reference pellets, and 12x12 pixels for all others, with pixel centers at the target separated by 0.5 mm steps). Since not all pellets move at the same rate, their positions need not be sampled uniformly in time. Economy of exposure dosage and data storage dictated differential pellet sample rates.

### 5.1.3.1. Pellet Sampling

Prior experience has shown that pellets in the vicinity of the lingual apex and blade (e.g., the pellet referred to as T1 in the database), are difficult to track successfully if they are not sampled at least 150 times per second. This rate is dictated less by signal bandwidth, which spectral analyses show less than 0-70 Hz (Figure 5.1), than by the simple tracking method employed by the XRMB system, which uses only the current and immediately previous position samples in predicting future raster positions. This tracking method, in spite
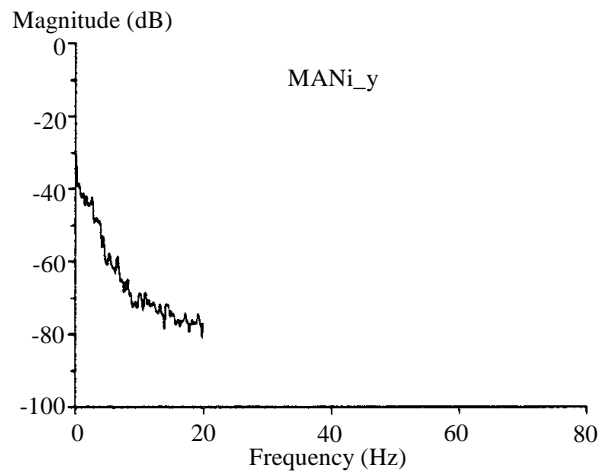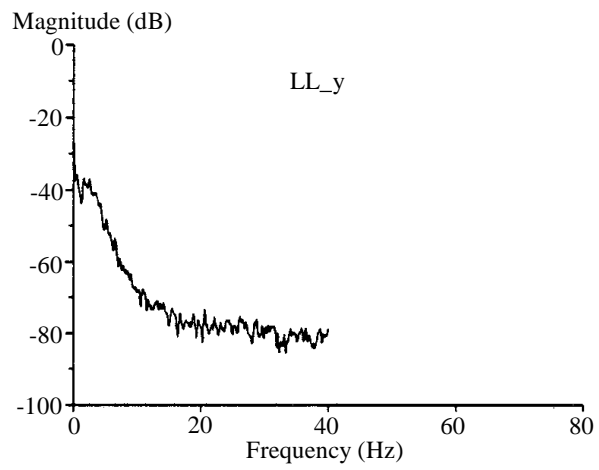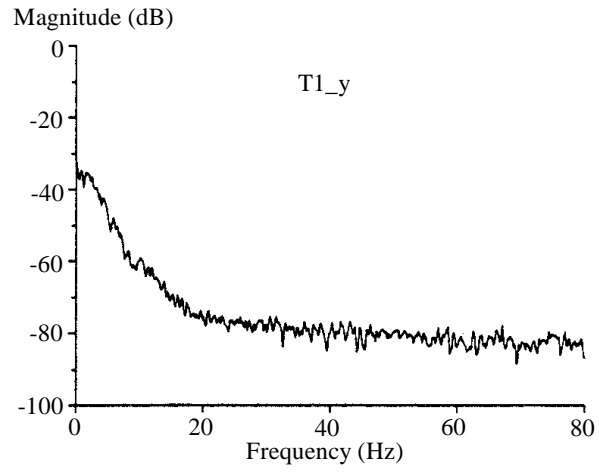
Figure 5.1: Spectra of y-dimension movements of T1, LL, and MANi pellets.

of the relatively wide 6x6 mm tracking rasters used for articulator pellets, is most susceptible to failure for fleshpoints that have rapid accelerations.  Apparently, that is a feature of the more ventral (anterior) aspect of the tongue during speech.

The starting point, then, for defining the array of sampling rates applied to pellet constellations for database speakers was to select an adequate rate for T1, at 160 samples/second. Rates for slower-sampled pellets were then specified according to a constraint of the channel-assignment-string of sequential-sampling control software.  This constraint required that the longer intersample intervals (dt's [Δt's]) associated with slower-sampled pellet rates be power-of-two multiples of the dt of the most-rapidly-sampled pellet.  Our usual custom was to sample pellets in real time according to the following schedule:

| Pellet type | Pellet name | N | Nominal sampling rate |
|---|---|---|---|
| reference | MAX(a) | 3 | 40 samples/second each |
| mandibular | MAN(a) | 2 | 40 |
| upper lip | UL | 1 | 40 |
| lower lip | LL | 1 | 80 |
| ventral tongue | T1 | 1 | 160 |
| mid-tongue | T2, T3 | 2 | 80 |
| dorsal tongue | T4 | 1 | 80 |

720 samples/second, aggregate

For some speakers with uncommonly quick and/or jerky movements of some pellets (most notably LL, T3 [dorsal mid-tongue], T4 [tongue dorsum], and MANi [mandibular incisor]), or, for those few speakers who did not finish with eleven pellets, the real-time sample-rate array was modified, though usually there was not much "headroom" available in the standard eleven-pellet constellation.  For example, 80 additional samples/s could be added (with further degradation in the system's ability to sample pellet positions in real time at the nominal rates) to LL or another pellet, or 40 samples/s to MANi, by reducing the rates for reference pellets to 20 samples/s.  Note that real-time rates were not preserved in fully post-processed data.  For several reasons (outlined in Section 6.2), raw pellet position histories were interpolated and resampled to an equal rate of 160 samples/s, with positions of all pellets defined at the same times.

### 5.1.3.2. System constraints on pellet sampling

Sample rates for any specific pellet are constrained by operational principles of the XRMB system related to target heating, and dependent upon beam voltage and current.  The maximum possible rate, for an electron beam voltage of 450kV and current of 1.75mA, is approximately 340 samples/second.  When an area of the tungsten target is struck by the electron beam more frequently than this, the target cannot be adequately cooled.  The risk is that the electron beam will burn through the target, resulting in a catastrophic system failure and a very expensive repair bill.  This operational constraint limits the maximal rate for any single pellet.  It also affects sample-rate constellations in another way:  spatially adjacent pellets may have intermittent **overlapping** trajectories and tracking rasters.  The obvious danger inherent to this condition, for

poorly-chosen combinations of sampling rates, is that the maximum aggregate sample rate for any single rastered area of the target may inadvertently be exceeded.

In all database sessions, beam voltage was 450 kV; in most sessions, current was 1.75 mA, but sometimes 1.5 or 2.0 mA. As a rule, we came to believe that higher beam currents, and their correspondingly richer distributions of photons per unit-image-plane area, were required for speakers whose pellets were difficult to track. A common response to poor tracking performance, especially for the T4 pellet (attached in the vicinity of the tongue dorsum), was to increase electron beam current by 0.25 mA. Entrance exposure is linearly proportional to beam current, and exposure estimates for database speakers were automatically indexed to relevant operating conditions during experiments.

Some raw pellet-position histories were not equal-time-interval when they were acquired. This feature of the data resulted from the channel-assignment string governing the real-time sampling. Consequently, selected raw signal streams required interpolation and re-sampling to make them equal-time interval (see Section 6.2). At the time of acquisition, time stamps to the nearest microsecond were stored with each pair of coordinates, for each sample of each pellet position. The time-stamp streams were subsequently incorporated in the re-sampling process, to maintain synchrony between the position time histories of pellets, and wide-band physiological channels. Within pellets, sampling jitter was systematic. Across pellets, jitter was greatest for the most rapidly-sampled pellet, T1, ranging +/- 12% with respect to the mean true intersample interval. Jitter was lower for the four pellets LL and T2-T4, nominally sampled 80 times/s, ranging maximally +/- 10% with respect to each pellet's mean intersample interval. There was no sampling jitter for raw position histories of the six slowly-sampled pellets MAX(a), UL, or MAN(a). In real time, these pellets were sampled at true equal time intervals.

It may also interest users to know that the actual dt's associated with raw pellet-position histories, even for pellets with true equal-time sampling, were not the nominal dt's corresponding to nominal (requested) sampling rates of 20, 40, 80, or 160 samples/s. The true mean rates were 18.8, 37.6, 75.2, and 150.4 samples/s, respectively. The discrepancy between nominal and true sampling frequencies resulted from interactions among three factors: (1) the particular constellation of eleven requested rates, and the associated serial array of samples and nulls required to traverse the channel assignment string; (2) the sizes of tracking rasters associated with each pellet (10x10 pixels, for each of three reference pellets, and 12x12 pixels for each of eight "articulator" pellets); and, (3) the ten-microsecond dwell time of the electron beam on each pixel area of the tungsten target. Together, these factors conspired to force the sampling hardware to acquire data slightly more slowly than the nominal requested rate. This performance limit inherent to the XRMB system was subsequently obscured in post-processing (see Section 6.2) which established equal-time and equal-rate intervals between pellet-position estimates, via interpolation and resampling, for all eleven pellet-position histories.

## 5.2. Experimental procedures

### 5.2.1. Pellet placement

A fundamental problem for the XRMB technique is the problem of knowing where, how, and how many pellets must be placed on a speaker in order to generate a meaningful view of the dance we call speech. This problem is shared by all equivalent transduction techniques that represent articulation in terms of point-parameters (i.e., movements of discrete fleshpoints and bony landmarks; Kent, 1972). For the point-parameterized view, expressed at the level of what we think of as articulator primitives (e.g., the lower lip, tongue blade, lower jaw, and the like), a crucial question is: "How much does one point count?" This problem is the converse of that posed by image data (e.g., high-speed cineradiography), where we must choose some seminal parameterization that best captures the essence of the data, and discards what is redundant and/or uninteresting. For both techniques, the hope is to extract a maximally economical characterization of all and only those articulatory features that are important and revealing.

For some structures, the answers to where, how, and how many pellets, are relatively clear and easily understood. This is true for both the mandible and the remainder of the skull, rigid bodies whose movements in free space can have six degrees of freedom at most. For other articulatory structures, the answers are less certain. There are no well-founded rules to guide pellet placement on the deformable tongue and lips, whose shapes are limited only by pliable, continuous boundary surfaces, constraints on compressibility, and perhaps some as-yet-unknown principles of functional compartmentalization. For the tongue and lips, we merely followed convention and the dictates of compromise. Pellet placement is a significant issue not resolvable with database materials alone, though these data will likely contribute to ongoing efforts to determine economical point-parameterization schemes for deformable articulators. For example, insights may come from synthesized speech based on pellet trajectories, compared to parallel recordings of the natural acoustic wave.

The same placement protocol was intended for all speakers, though in some cases, procedures were modified according to speaker anatomy, function, and/or integrity of the initial adhesive bond. Pellets were glued to the tongue using Ketac, and to all other surfaces using Isodent (commercially-available dental adhesives), and then anchored by light threads taped to the skin surface of the cheeks and face. Relatively few constraints affected pellet placement. The most notable constraints were a speaker's sensitivity and gag reflex; inter-pellet spacing (generally never less than 1 cm, on center, and somewhat more along the tongue); the number of pellets that could be successfully tracked at one time (limited by aggregate sampling rate, and the manner in which that aggregate could be distributed among pellets, relative to structure-specific accelerations); and, speaker-specific patterns of movement producing unacceptably frequent mistracking, due to overlapping trajectories of two or more pellets.

The usual pellet constellation, intended for each database speaker, included eleven pellets, placed generally according to the scheme shown in Figure 5.2. The name associated with each pellet accompanies its placement description below, and is shown in parentheses.

Figure 5.2: Approximate pellet placement locations
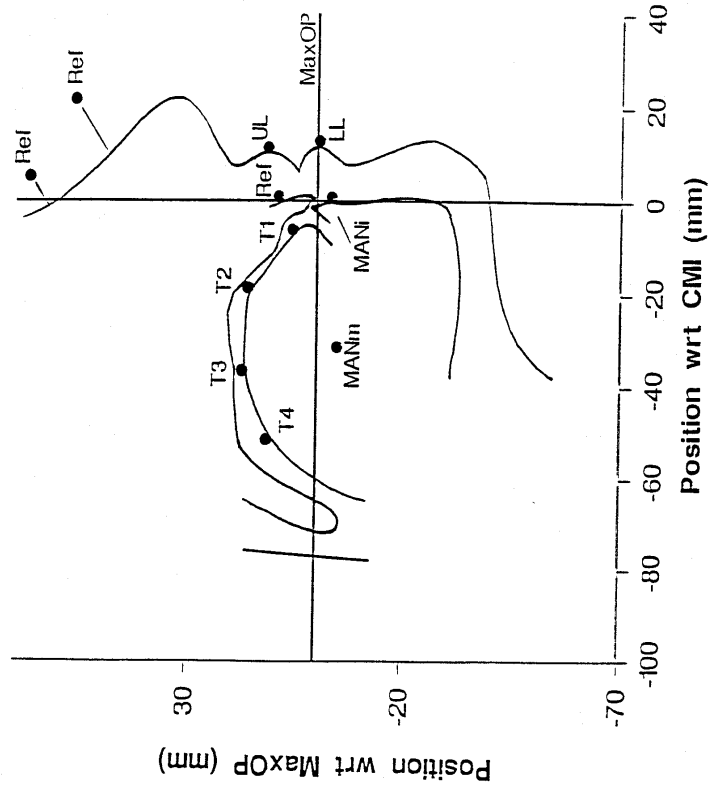
*5.2.1.1. Reference pellets*

A reference (or fiducial) triangle was formed by three pellets attached to the speaker's head: one (MAXn) placed on a short stand-off post glued along the bridge of the nose; a second (MAXi) glued to the buccal surface of the maxillary incisors, in the pocket formed by the central diastema and the enamel-gingival border; and a third (MAXg) also placed on a stand-off, glued lower along the nosebridge (for speaker numbers greater than JW17), or attached to an arm projecting from a snug-fitting pair of eyeglass frames. Two of the reference pellets, highest on the nose bridge and at the incisors, were routinely used to establish a floating two-dimensional, cranial-based coordinate system within which the positions of other pellets could be described. Defining the coordinate system in this way allowed complete removal of head motion from the motions of the remaining "articulator" pellets, as long as head motion was simple, involving only pitching rotation (about axes normal to the midsagittal plane), and/or translation relative to axes lying in the midsagittal plane. However, the head was never restrained in any experiment, and more complex head movements were possible. Accordingly, the triangle of reference pellets was originally intended to be used according to a single-view, calculating procedure developed for tracking head motion in three dimensions (Westbury, 1991). Unfortunately, after many speakers had been recorded, we concluded that the inter-pellet distances for the reference triangle could not be measured with sufficient accuracy required by the calculating method, and that the triangle itself was too "degenerate" (i.e., its vertices were essentially coplanar, and at least one angle was too acute) to permit unambiguous, three-dimensional reconstructions of head position. Nevertheless, we continued to track three reference pellets in all sessions, and eventually discovered that the third, superfluous pellet proved useful in those records where one of the two reference pellets typically used to define the cranial coordinate system had been mistracked. In records of this type, the third reference pellet allowed recovery of data that would otherwise have been lost to tracking failure.

The placement of reference pellets was variable across speakers, for reasons that are all too easy to imagine. For example, no two people have noses the same size and shape. Consequently, the reference pellets themselves could not be used to standardize the anatomically-defined coordinate system ultimately imposed on each speaker's data. A special protocol, using a calibrated biteplate (see Section 5.2.2.1, below; also, cf., Westbury, 1991), was established for that purpose. In short, the standard coordinate system, defined individually for each speaker (see Section 5.2.2.1, below), had its origin centered in the diastema formed by the exposed tips of the central maxillary incisors. Its x-axis corresponded to the intersection of the midsagittal plane and a second plane, referred to as the maxillary occlusal plane (hereafter, MaxOP), given by the tips of the central incisors and at least two other maxillary teeth on opposite sides of the mouth. The midsagittal plane was itself assumed to be normal to the MaxOP, containing the line passing through the central maxillary diastema and points midway between matched cusp tips on opposite sides of the maxillary arch. The y axis was then normal to the MaxOP, and intersected that plane at the origin, with the +x (ventral, or anterior) and +y (rostral, or superior) directions out of the mouth, and toward the roof of the mouth, respectively.

This standard anatomically-based reference frame, common to all speakers, is one of the underlying strengths of the XRMB Speech Production Database. The benefits of the framework are several, and include elimination of coordinate-system bias in data description and

interpretation; quick communication of data, due to shared familiarity with the representational scheme; and intuitively satisfying direction senses, where "up" (rostral, or superior) may be thought of as toward the top of the head (along lines normal to MaxOP), and "forward" (ventral, or anterior) out of the mouth and toward the face front (along lines parallel to the intersection of the midsagittal and maxillary occlusal planes).

### 5.2.1.2. Mandibular pellets

Two pellets were routinely attached to the mandible: one (MANi) glued to the buccal surface of the central incisors, analogously to the maxillary incisal pellet, in the pocket formed by the central diastema and the enamel-gingival border; and, a second (MANm) glued in the vicinity of the juncture between the first and second mandibular molars, typically on the speaker's left side, sometimes as rostral as the enamel-gingival border, but more commonly glued to the gingiva itself.

The mandible, like the remainder of the skull, is a rigid body. As long as the motions of its two halves mirror one another about the mid-sagittal plane, a complete and general account of its motion will be given by the trajectories of two points. Modest amounts of data in the literature consistently show that treating the mandible in this fashion is both appropriate and necessary for speech, since there are measurable and systematically-independent amounts of sagittal-plane rotation and translation associated with articulatory behavior, and essentially negligible amounts of transverse and coronal plane rotations and translations (Gibbs & Messerman, 1971; Vatikiotis-Bateson & Ostry, 1993).

For some speakers, the MANm trajectory frequently intersected the trajectories of some lingual pellets, and data were then lost to mistracking. Pellet repositioning sometimes helped, but detachment was necessary in other cases. For all speakers of this latter type except one, the offending MANm pellet was removed, and the experiment conducted without it. Point-parameterized kinematic data for the tongue are much more difficult to obtain than for the mandible. Among those few speakers where MANm removal was necessary, it is no longer possible to properly de-couple the lingual and lower-lip fleshpoints from the mandible, taking into account both mandibular translation and rotation.

### 5.2.1.3. Lingual pellets

Four pellets were attached along the longitudinal sulcus of each speaker's tongue. The most ventral of these was typically placed in the vicinity of the so-called tongue "blade," roughly 10 mm posterior to the apex of the extended tongue. The most dorsal was placed about 60 mm posterior to the apex, as far back as the speaker would tolerate without gagging, but always ventral to the circumvallate papillae. Two medial pellets were placed so that the distance between the front and rear-most pellets was divided into three roughly equal segments.
The optimum number of pellets necessary to represent tongue motion during speech is unknown. On the basis of a small set of XRMB image data for static vowel postures, involving a string of closely-spaced pellets laid along the tongue length, Lindau et al. (1988) suggested that only three pellets are necessary to accurately predict the midline contour extending from near the teeth, rearward to about the tongue root (ca. the lingual valleculae). However, Martin's (1991)

microbeam study of swallowing indicated that at least five lingual pellets may not be too many. We chose to use four whenever possible, a workable compromise between a constraint on minimum interpellet spacing and the desire to minimize gluing time. Exceptions to the four-pellet norm, for some speakers and/or records resulted from over-lapping pellet trajectories (e.g., overlapping tongue and MANm pellets), or broken adhesive bonds. In general, a lingual pellet that came loose during an experiment was not replaced if more than 15% of the tasks had been recorded. Too much uncertainty was associated with duplicating the initial placement. In contrast, reference, mandibular, and labial pellets that became detached were routinely replaced, because constancy of placement was either irrelevant (for the rigid bodies), or easier to satisfy (for the lips).

Table 5.1 provides a numerical summary of lingual pellet placements, in millimeters posterior to the apex, measured along the surface of the extended tongue and averaged across speakers. This information is graphically summarized in Figure 5.3. Specific placement locations for each speaker are available in a text file accompanying the Database.

Table 5.1:  Pellet placement (mm) from tongue apex

| Location | Pellet Name | Mean | S.D. | Range |
|---|---|---|---|---|
| ventral | T1 | 8.5 | 1.07 | 7-12 |
| mid-ventral | T2 | 25.2 | 2.44 | 19-31 |
| mid-dorsal | T3 | 43.8 | 3.49 | 33-56 |
| dorsal | T4 | 60.1 | 4.14 | 45-73 |

Customary spacings between adjacent tongue pellets, averaged across speakers, can be derived from placement measures. Spacing measures in Table 5.2 show, at least on average, that pellets were relatively evenly distributed along the length of the tongue.

Table 5.2:  Spacing between pellets

| Pellet Pair | Mean Spacing | S.D. | Range |
|---|---|---|---|
| T1:T2 | 16.7 | 2.02 | 12-22 |
| T2:T3 | 18.6 | 2.72 | 12-27 |
| T3:T4 | 16.3 | 2.81 | 11-26 |

Overall, there were no significant gender-related differences in placement or spacing of lingual pellets. The speaker with closest lingual-pellet spacing was JW52 (female: @7,21,33 & 45 mm). She may have had a robust gag reflex, but stone models of her teeth show neither a small maxillary arch nor small palatal vault. The speaker with widest spacing was JW42 (male: @9,31,56 & 73 mm): no monster, according to his models; probably just insensitive.

Figure 5.3: Means and standard deviations for lingual pellet placements

41
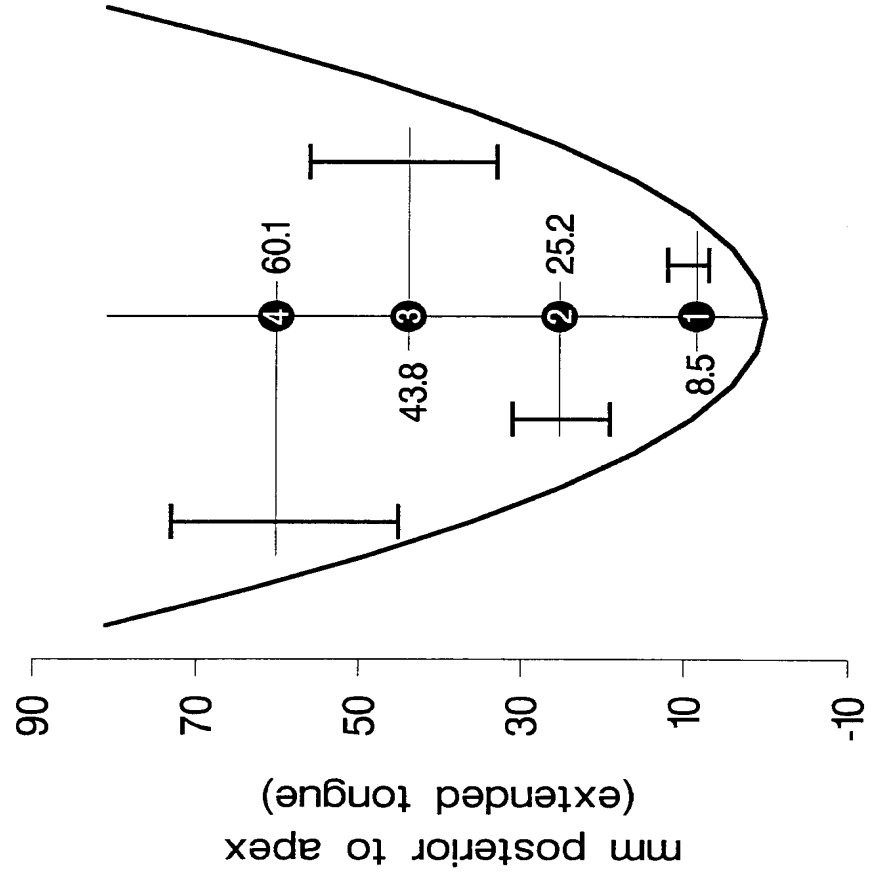
### 5.2.1.4. Labial pellets

One pellet each was attached to the upper (UL) and lower lip (LL), glued to the external surface at the vermillion border. The lips are relatively more accessible than the tongue for movement transduction and analysis, but are probably no better understood. In particular, relationships between (single) point-parameterized representations of labial function, like those supplied by the XRMB technique, and speech-function concerns with lip aperture, rounding, protrusion, and the like, are poorly documented in the relevant literature. Certainly our choice of one midline pellet per lip provides an impoverished view of labial function, though this view is likely no worse than those available from strain-gauge, photo-optical, and electromagnetic techniques used by others in the field.

### 5.2.1.5. Pellet size and speech sound production effects

Lingual and labial pellets were usually 2.5 mm in diameter, and mandibular and reference pellets were typically 3 mm in diameter. Pellets attached to the tongue, in particular, were embedded in a shallow mound of adhesive, perhaps 5 mm in diameter, and at least half as deep as the pellet diameter itself. It is therefore no surprise, from a qualitative point of view, that the pellets in many speakers' mouths had a perceptible effect on their speech sound production, at least for a brief period after initial pellet attachment. Some speakers adapted to pellets better, and more quickly, than others. For many, the difference between speech with and without pellets was largely unnoticeable, particularly for records late in the inventory. For a few, the effect of pellets on sound production was more robust, though not always in expected ways. We remember one speaker who passed the screening assessment with no problem, yet appeared for data collection unexpectedly showing mild but noticeable fricative distortions. The distortions were apparently "cured" by pellet attachments to his tongue and teeth. That speaker notwithstanding, the perceivable effects of pellets were variable across speakers, and within speakers over time. However, as of this writing, no fine-grained listening studies have been performed to assess perceivable acoustic effects of pellets. Partly to allow such assessment, approximately 20% of the task inventory was pre-recorded for each speaker, without pellets. These data were digitized and archived using the same system used to record the pellet-tracking data, and can be analyzed to develop a measure of the degree of sound distortion associated with pellets. The distortion, whatever its extent, is merely a necessary cost of the technique, and does not preclude matched task-and-condition comparisons within and across speakers.

### 5.2.2. Special microbeam records

Most of a recording session for each Database speaker involved a series of trials in which pellet positions were tracked during performance of a designated series of speech and oral motor tasks. Four types of special records, required for calibration, definition of certain representational conventions, and system initialization, were also collected and interspersed among the task-tracking records.

*5.2.2.1. Biteplate record*

Following a procedure described elsewhere (Westbury, 1991), a calibrated, triangular arrangement of three pellets, illustrated in Figure 5.4, was fabricated for each subject. This pellet triangle, an aluminum plate on which it was mounted, and an attached silicone-base impression of the subject's maxillary teeth, together comprised a calibration *biteplate*. The shallow impression of the subject's maxillary teeth were mounted onto the 3x58x116 mm aluminum plate (representing the MaxOP), such that indentations made by the central incisors and at least two teeth on opposite sides of the maxillary arch cut through to the aluminum surface. The speaker's midsagittal plane was assumed to be normal to the plate, and to contain the line passing between the central incisors and a series of midpoints between matched cusp-tip locations in the impression. Two gold pellets were embedded in the biteplate, both on the midsagittal line: one immediately beneath the central incisors, to represent the origin of the cranial reference plane in which movements of articulator fleshpoints were subsequently described; and the second, ventral to the incisors, defining the +x-direction of the midsagittal plane. A third pellet was mounted off the biteplate surface, a known distance lateral to the speaker's midsagittal line.

At the outset of the recording session, positions of the three biteplate pellets, and the three reference pellets attached to the speaker's head, were tracked concurrently for a five-second interval, at 60 samples/s in an enhanced resolution mode, while the biteplate was held in tight contact with the maxillary teeth (Figure 5.5). Average image-plane coordinates of all six pellets were then calculated, in part to establish the positions of cranial reference pellets tracked during all other records, relative to the standard MaxOP coordinate system eventually imposed on articulatory data. The relative positions of biteplate and reference pellets were eventually supplied as input parameters to post-processing routines (see Section 6.3) which operated sample by sample to remove sagittal-plane motions of the head from position histories of all remaining articulatory pellets.

The mean biteplate-pellet coordinates determined from this record were also supplied as inputs to a different calculating routine, along with appropriate specifications for pinhole-to-image-plane and inter-pellet distances. Resulting approximations of three-space coordinates for the three biteplate pellets were used, in turn, to determine a local, three-dimensional, orthogonal axis system for the head, with the origin at the mid-sagittal pellet immediately below the central maxillary incisors, two axes lying in the MaxOP, and the third normal to the MaxOP. The local, cranial axes were finally used to determine a set of Euler angles and a corresponding rotation matrix (cf., Goldstein, 1980) that together indicated the relative orientation of the three-dimensional cranial and machine-based coordinate systems.

Broadly, this procedure was used to evaluate each speaker's head position, in three dimensional space, relative to the position assumed by microbeam system operation (see Section 6.4). An evaluation of this type was necessary because, as usual for XRMB experiments, head position was not constrained for any database speaker. Results from calculations provided estimates of measurement error arising from improper head positioning, and indicated that data for all speakers were within 3% of being true-to-life.
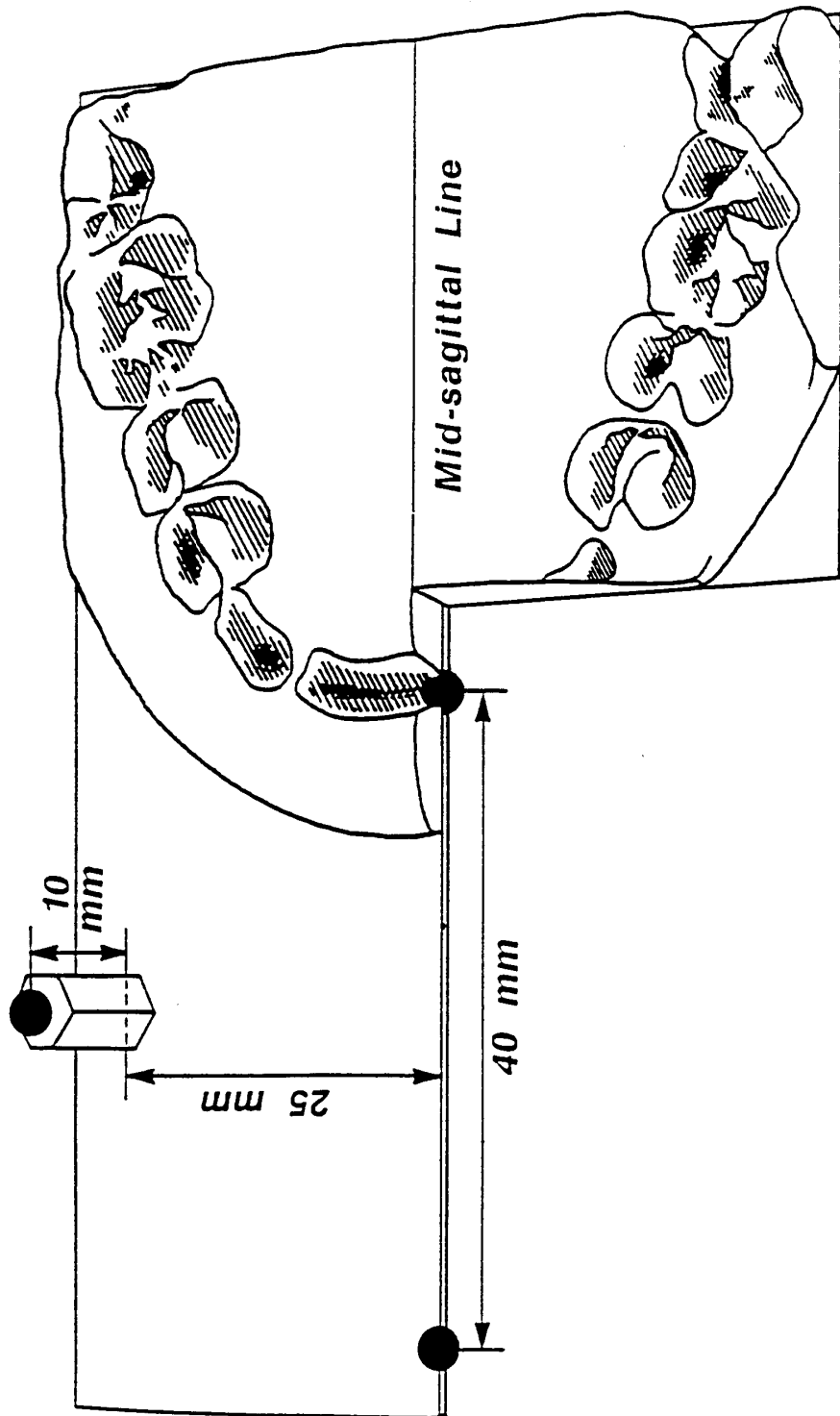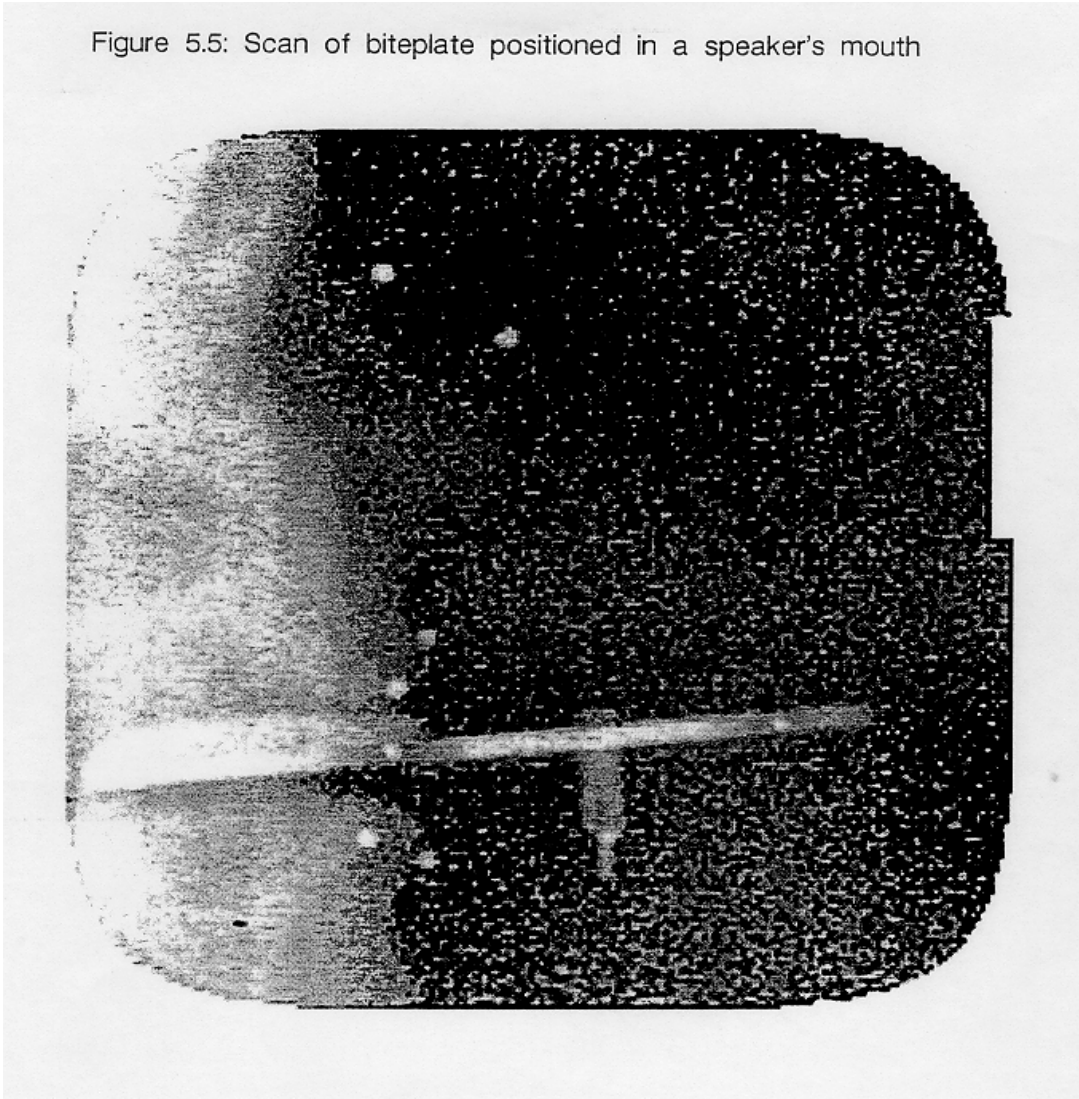
Figure 5.4: Schematic of the biteplate

Figure 5.5: Scan of biteplate positioned in a speaker's mouth

*5.2.2.2. Icon*

At the end of some speakers' recording sessions, perhaps 1 in 6, the static positions of 5-10 pellets in a cross-shaped calibration array (icon) were tracked for several seconds.  Estimates of relative pellet positions derived from this record were used to document and verify system gain, in both dimensions of the image field.  The XRMB system is a complex device, and sometimes requires physical maintenance.  Recurring maintenance operations, over the two-year period when data were acquired, sometimes had significant effects on beam deflection.  System gains were routinely restored after such operations, following evaluation of icon-tracking records.

*5.2.2.3. Initialization scans*

The XRMB system is massive, and its position within the laboratory is fixed.  In contrast, a speaker's position, and especially the relative positions of pellets within the system image field, are not fixed.  At the beginning of every record, the system must be told where to look for pellets before tracking can be initiated.  These specifications for expected first pellet positions are initially generated from a procedure referred to as an *initialization scan*, involving a rapid, 400 ms sweep of the X-ray beam over the central 15x15 cm portion of the image field.  The resulting scan image, derived from a 10 microsecond-per-unit-area exposure, resembles a conventional, albeit fuzzy, radiographic image that can then be displayed to a system operator on a graphics monitor.  The operator visually estimates pellet locations, and interactively enters them into a command file that must be read by computers responsible for first aiming the beam at the beginning of the following record.  When the system is triggered to track, local searches of predefined pellet positions are initiated.  If all pellets can then be found from these automated searches, full tracking is initiated, and a cue tone is generated, signaling the speaker to begin performing.

As long as the speaker remains still, between the time of the initialization scan and the trigger for the following record (typically about 45 seconds), successful tracking will occur.  Speakers that wiggled a lot, between records, often required frequent initialization scans, sometimes before every third or fourth record.  Other speakers, capable of sitting very still for long periods, were able to run as many as 20-30 tracking records without new initialization scans.

Initialization scans are routine because they are prerequisites for tracking, both for the first record for each speaker, and for all records preceded by significant movement. Interestingly, they also provide valuable information concerning speakers' anatomy (e.g., the posterior pharyngeal wall outline described below; and, areas of dense bone, or fillings, that sometimes make tracking difficult).

*5.2.2.4. Vocal tract boundary outlines*

It is tiresome to hear the left-field fence at Boston's Fenway Park always called the Green Monster.  Cliches wear patience like Spring wears hope for the Sox.  Cliche or not, no one can watch the poor Bean'ers play at home and forget the effect the great wall has on their game.  Hitters dance from its seduction; left fielders must play its kiss.  The Monster tunes every game

to be something it would not be any time and any place else.  Daily lines of Sox 9-8, O's 6-5, Yanks 12-7, crazy numbers anywhere else, whine always at the Fens.

As in baseball, perhaps so it is in speech.  Every speaker's mouth is bounded by walls with special shapes and sizes, and we guess that these mold the speaker's movements as powerfully as the Fenway Monster molds its game.  Knowing the outlines of these walls will surely tell us something about why positions of the tongue are distributed as they are during speech.  The outlines are useful for other reasons, as well.  Certainly they add a sense of realism to records of articulatory movement.  The tongue routinely contacts the palate for some consonants.  Seeing it do so in the data streams, at about the times and places we think it ought to, is reassuring. Knowing the relative positions of tongue and boundary also means that we have a way of estimating at least some portion of the vocal-tract area function and its associated acoustic transfer function.  Finally, the boundary outlines afford useful information about size of the space where each speaker must work, providing estimates, for example, of depth and length of the oral cavity.

For each database speaker, these walls are given in the form of a midsagittal outline of the palatal vault, and a line segment representing some short stretch of the posterior pharyngeal wall (ppw), as illustrated in Figure 5.6.

*5.2.2.4.1. Palatal outline*

At least one of two methods was used to define a palatal outline for each speaker.  For the majority of speakers, the palatal outline was determined from an averaging scan (ca. 25-100 sweeps) of the calibration biteplate (and its associated reference-pellet triangle),  mounted atop a stone model (plaster dental cast) of the maxillary dental arch, with a string of gold pellets laid along the midline of the palatal vault.  Static images of this type (see Figure 5.7) were corrected for image plane distance and target curvature, and displayed using a software "Locator Tool." Pellet-center locations were then visually identified and hand-marked using an interactive cursor. Subsequently, the palatal curve was approximated by a piece-wise continuous function formed by line segments linking adjacent pellet centers.  In cases where pellets were widely spaced, separated by 3 mm or more, attempts were sometimes made to lessen the effect of coarse spatial sampling, and improve the resulting ragged outline, by picking from the image other (non-pellet) points along appropriate regions of the palatal curve.

The palate outline derived from this procedure can be aligned with the pellet trajectories from tracking records, since both the pellet tracks and outline were expressed relative to the local coordinate system defined by the calibration biteplate.  A record holding a palatal curve of this type has been appended to the dataset of each subject for whom appropriate materials were available, and can be displayed concurrently with pellet trajectories for evaluation.

For a few speakers, palatal outlines were also generated by tracking the motion of a tracing pellet, manipulated by the experimenter or subject, drawn slowly along the palate midline.  Thus, for these speakers, two outlines may exist.  Fiducial pellets attached to the head, defining the standard cranial coordinate system, were tracked as the trace was performed, thus allowing spatial alignment of articulator pellet tracks and the palate outline tracing, through

reference to the fiducial markers.  For a very few speakers, there is only a manual-trace palatal outline.  For these speakers, approximating the outline from the stone model was either impossible or impractical.  JW52 is one such speaker, for whom the quality of the maxillary stone model was unacceptably poor, owing to a lump of plaster extending over much of the palatal vault.

Palatal outlines derived from either procedure span an interval of approximately 5-55 mm dorsal to the CMI.  The ventral-most point in each outline represents either the position of the first pellet in the palatal chain, or the location along the palate where the speaker stopped tracing forward.  Similarly, the dorsal-most outline point represents either the position of the last pellet in the palatal chain, or the location where the speaker stopped tracing rearward.

It is possible to extend the palatal outline toward the pharyngeal wall, beyond the last pellet-chain position or dorsal-most trace, by noting the clustering of extreme rostral positions of T3 and T4 pellets when their trajectories traverse regions not bounded above by the original outline (i.e., regions dorsal to the dorsal-most point of the palate trace).  An example of palate extension is illustrated in Figure 5.8.  Extreme pellet positions may represent the palatal boundary, especially during sounds requiring dorso-rostral palatal contact (e.g., [k] & [g]).  But, it's important to remember that extending the palatal outline in this fashion, to span a region where the palate is soft, is difficult and uncertain, because it is in this area that the boundary position itself varies.  If this portion of the outline is estimated on the basis of *oral stop* extrema, for example, when the soft palate is likely to be high, it may appear that the tongue does not contact the palate during velar *nasals* (where the soft palate must be low).  This inference is surely false, and exemplifies the problem of approximating a movable boundary with a single estimate, as if it were immovable.  Knowing pellet positions with respect to the soft, dorsal part of the palatal curve, where the palate is mobile, is inherently problematic unless we have some independent measure of where that portion of the palate happens to be.

Ideally, there should be good agreement between the two versions of the palatal outline, generated from scan images and tracing, for speakers who have them both.  After all, there was only one curve for any speaker, and how we traced it shouldn't matter.  In fact, poor agreement was  uncommon between the traced and scan-generated versions of the curve, for speakers who had them both.  The two curves lay closely together, at least ventrally.  Section 6.6 contains additional details on palate generation and its alignment with pellet data.

At least one good approximation of the midsagittal outline for each speaker's palatal vault is included as a special record in each dataset.

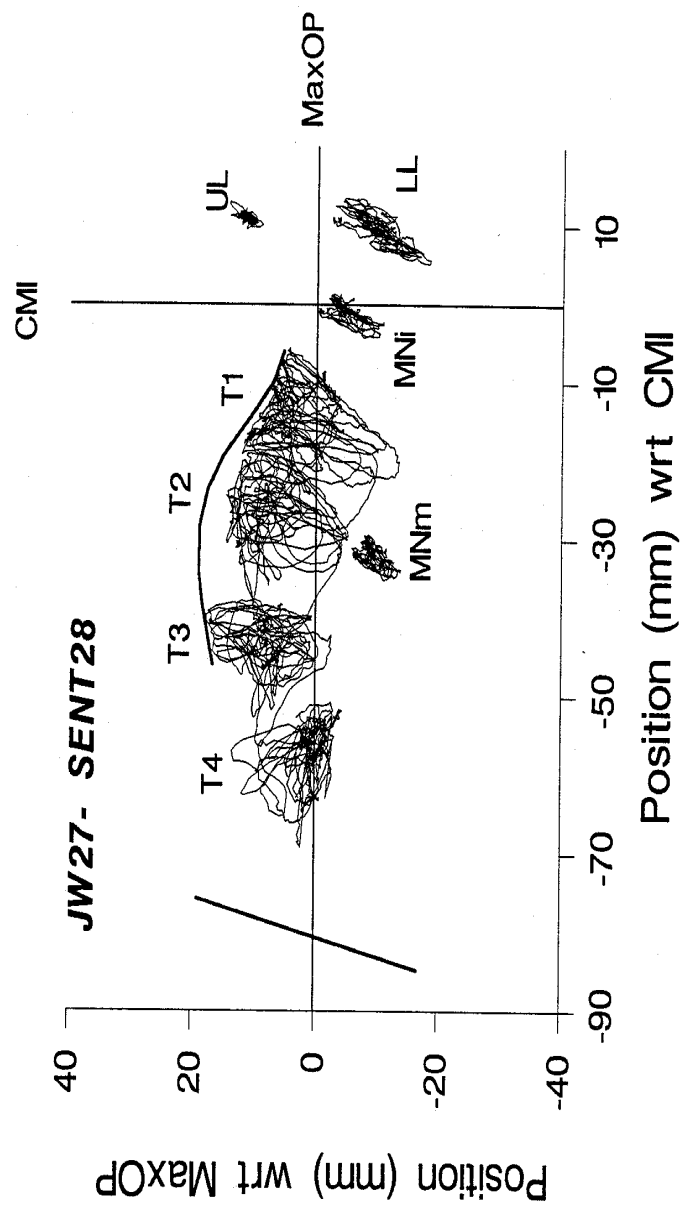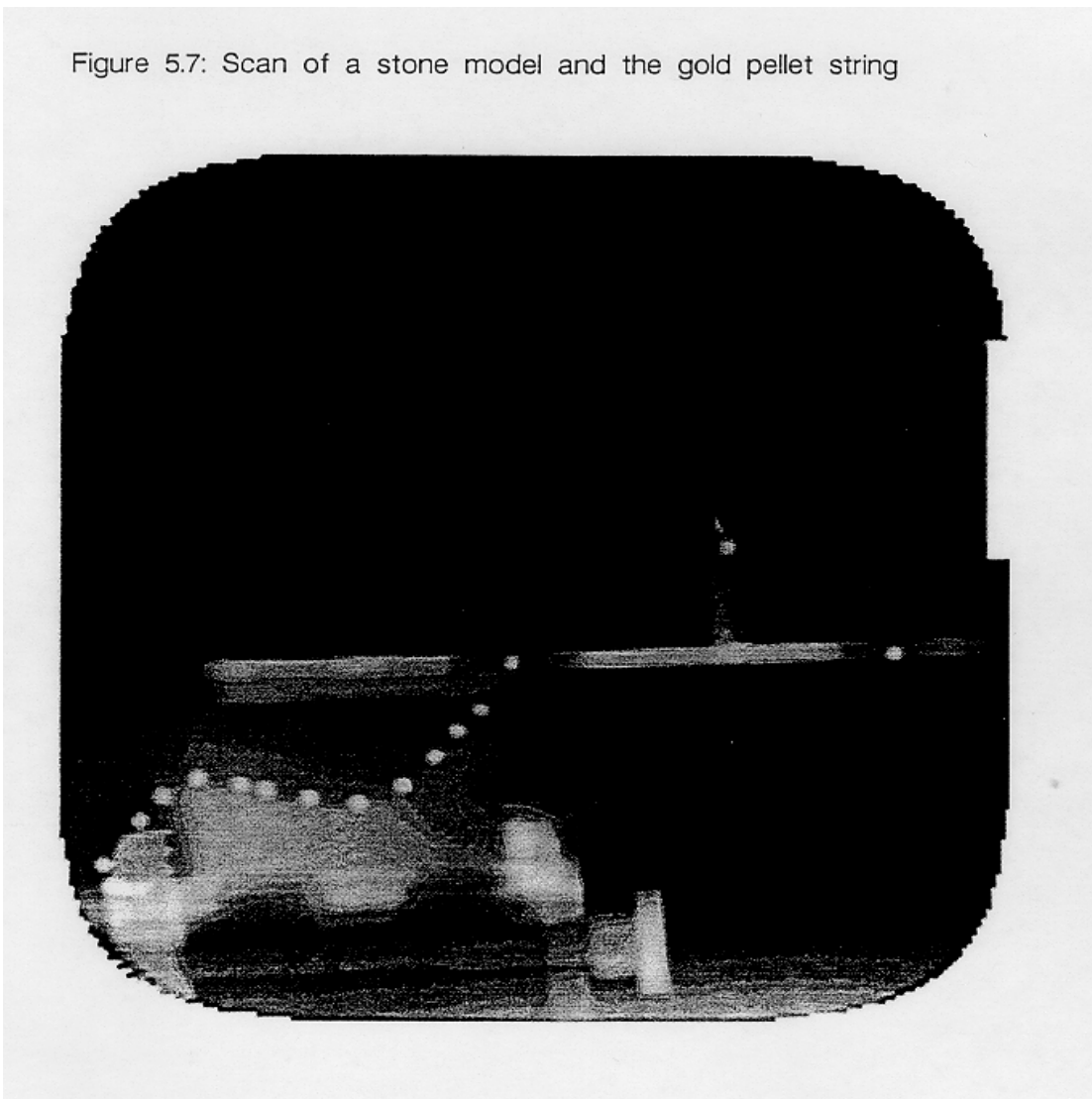Figure 5.6: Midsagittal plane trajectories with both palate and pharyngeal outlines

Figure 5.7: Scan of a stone model and the gold pellet string

*5.2.2.4.2. Posterior pharyngeal wall outline*

An approximation of the location and orientation of each speaker's posterior pharyngeal wall (ppw) was derived following a procedure similar to that used for the scan-generated piecewise continuous approximation of the palatal curve.  An initialization scan preceding the first tracking record, illustrated in Figure 5.9, was used to locate two arbitrary points along the ppw contour.  The straight line segment connecting them forms a coarse representation of a portion of the ppw. In a few cases, if the palatal outline in the initialization scan preceding the first tracking record was out of field, or insufficiently clear, a later initialization scan was used.

The vocal tract images captured in initialization scans were never sharp.  The special ppw records derived from them provide useful, albeit coarse, approximations of the position and orientation of the dorsal-most surface bounding the pharynx.

Figure 5.8: Extension of palate based on T4 extrema
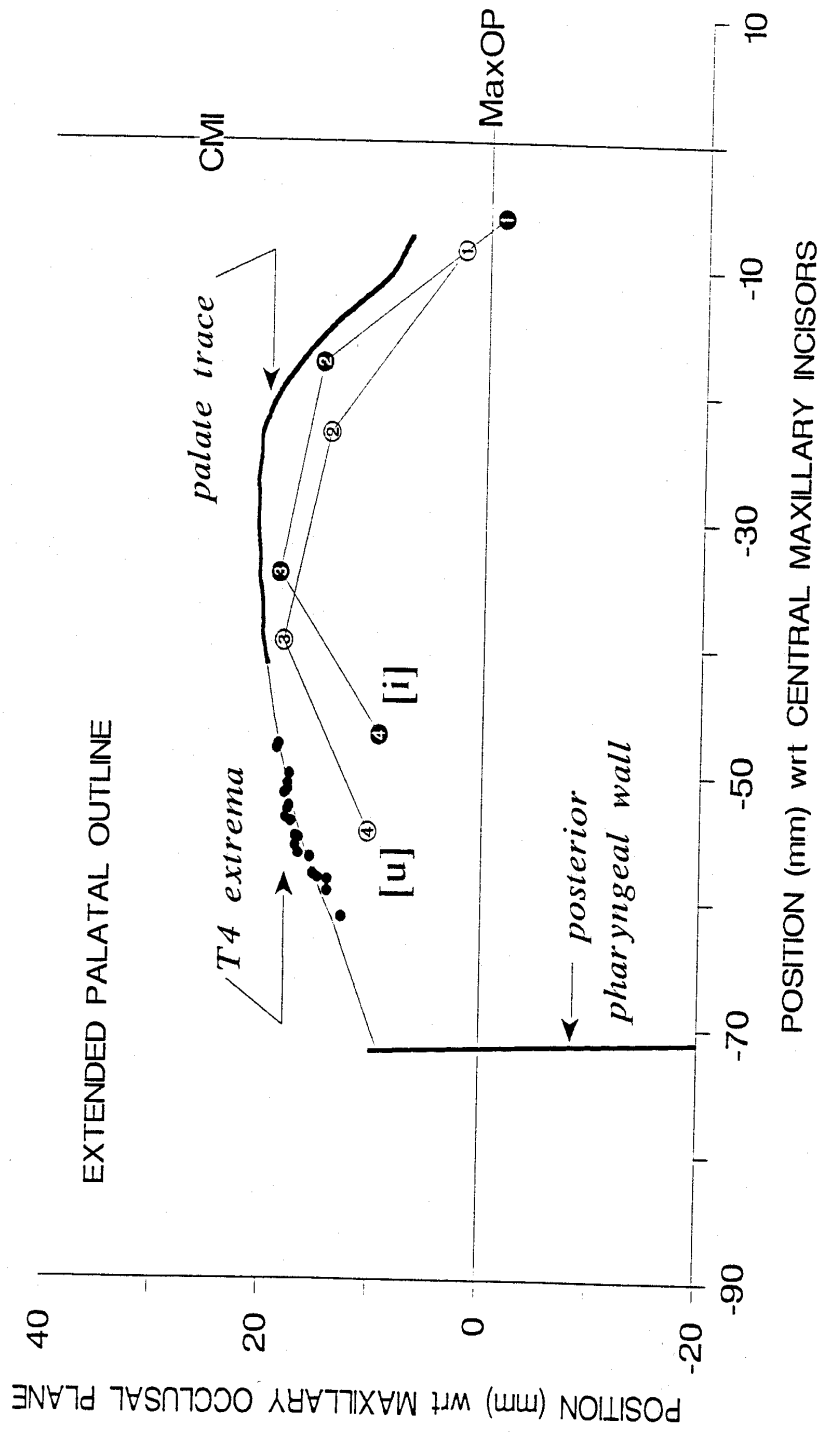
Figure 5.9: Midsagittal plane scan with line representing the PPW
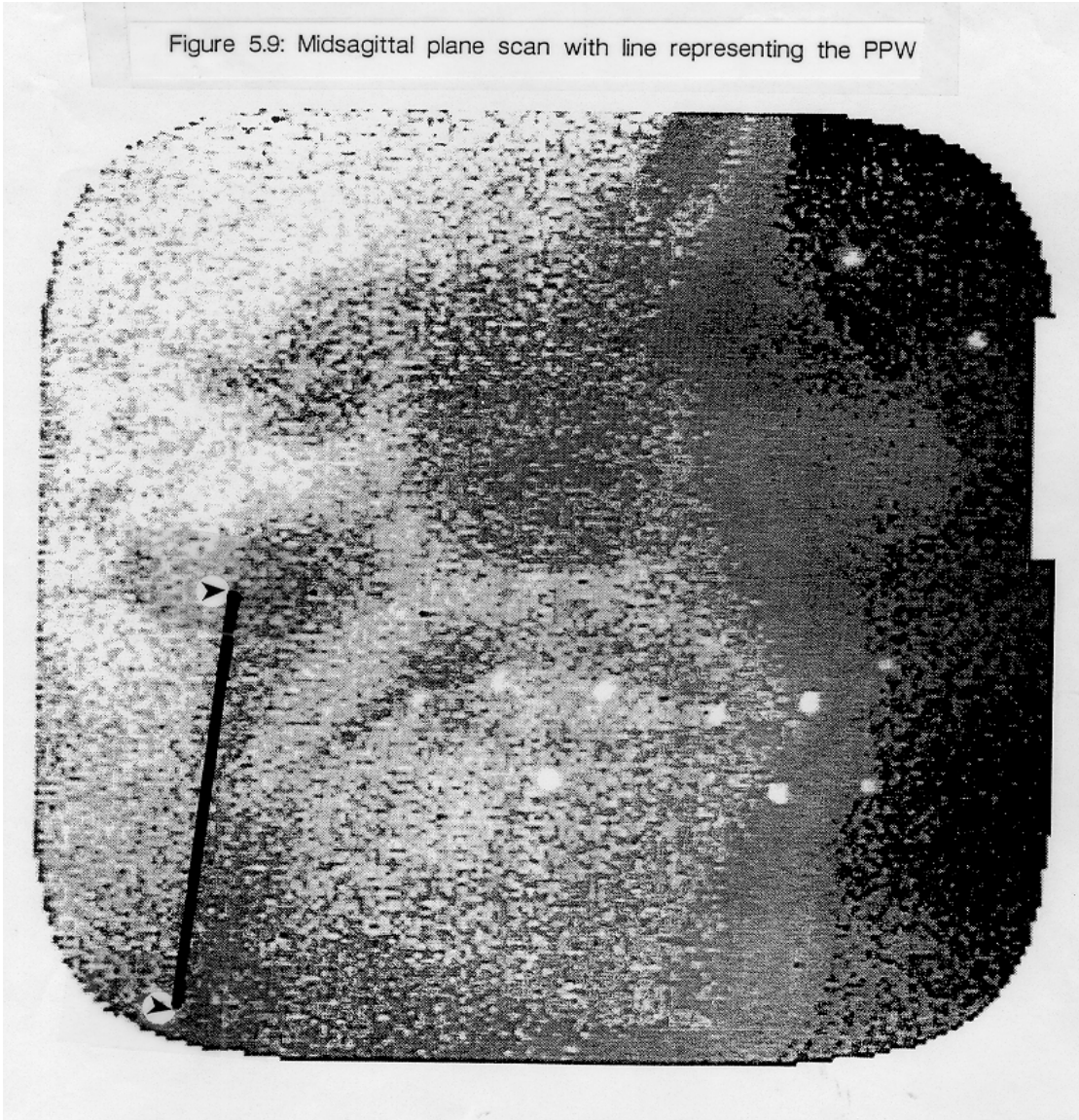
## 5.2.3. Speaker positioning

Throughout each recording session, speakers were seated comfortably in a cushioned, motorized, reclining examination chair, most often with their lower legs and feet propped and slightly raised. The head was supported from the rear by a headrest, but otherwise unrestrained. Even soft restraints become intolerable when they are maintained for 2 to 3-hour intervals.

Speakers were encouraged to remain as still as possible during the recording session, as they progressed through the sequence of 118 records in the task inventory. Sitting still was important for two reasons. Foremost, the constellation of pellets attached to the speaker's head had to be kept within the 20x20 cm field of the XRMB system. At the same time, each pellet had to be kept within the 1 square-cm area of the image field corresponding to its respective search raster. Whenever these conditions were not met, tracking could not be initiated, and new initialization scans were required. Keeping the head still, and positioned so that the midsagittal plane closely approximated the system image plane, was also important because the accuracy of data returned by the system is related to this condition (Westbury, 1991). When the two planes do not correspond, a scale error directly proportional to the discrepancy is introduced.

The system operator monitored position of the speaker's head during the recording session by viewing lateral and frontal video images as they were recorded (see Section 5.1.2.). Speakers also monitored their own position by viewing a mirror reflection of a low-intensity laser beam projected onto their forehead. At the beginning of each successive record, speakers were instructed to visually match the position of the laser to a small target pasted on the forehead, and then to lightly intercuspate the maxillary and mandibular teeth and rest the tongue in the floor of the mouth. In this way, speakers could reliably re-approximate the starting posture required for tracking initiation. Once tracking began, speakers were free to move the head in ways that were natural, as they performed the task. The fact that such movements routinely occur was the primary reason that fiducial pellets attached to the head had to be tracked.

Recording breaks were permitted, whenever any speaker requested them. A stretch break, roughly half-way through each session, allowed speakers and experimenters to walk about and relax. Fatigue was a significant opponent in most sessions. Some speakers found it increasingly difficult to maintain a steady position as they progressed through the task inventory. For these speakers, initialization scans, and the delays associated with them, were repeated with increasing frequency.

Head position, during each record for each speaker, was evaluated off-line, by several methods, in both two and three dimensions. The results of some of those evaluations were important for post-processing the pellet position histories, and are described in detail elsewhere ( Section 6.4).

## 5.2.4. Speaker-experimenter interaction

An experimenter was present during all recording sessions, partly to monitor the acceptability of task performances by speakers. When speakers mis-read or mis-articulated material particularly badly, the experimenter might request that the task be re-recorded. When

speakers were unsure about how tasks should be performed,  not a rare event for naive speakers, the experimenter would provide clarification or task modeling.

As a rule, however, interactions between speakers and the experiment team were minimized, partly because they tended to draw speakers' attention away from the problem of head positioning.  Also, our intent was to obtain from speakers as natural a performance as possible, albeit under unnatural circumstances.  This is not difficult for familiar material (real words, sentences, and the like) of the type that dominate the task inventory.  Interactive coaching was usually necessary only for "nonsense" utterance records, and for close adherence to speaking-rate constraints imposed by fixed record-lengths.  Interruptions for repeated records, except in the most egregious or critical of cases, were avoided to minimize radiation dosage.  The XRMB technique, by design, limits speakers' exposure to ionizing radiation by sampling only those areas where pellets are expected, and by sampling only as frequently as necessary for successful tracking.  Multiple requests for repeated tasks, driven by misguided dreams of perfect performance, run counter to the philosophy that risk should be minimized.

### 5.2.5. Task prompting

All tasks and/or task descriptions were displayed under experimenter control on a graphics terminal placed no more that 1 m directly in front of each speaker.  Speakers were instructed to read, repeat, or perform each prompted task, following an audible cue tone.  The starting cue was itself triggered by successful first recognition and location of all pellets in the tracking constellation.  First recognition depended, in turn, on speakers' ability to accurately achieve (or return to) the "ready" posture required by the XRMB method.  When speakers believed they had (re-)established the ready posture, and were certain of the task to be performed, they cued the system operator to trigger the x-ray generator and pellet search routines.

All speech in the production database represents fluent reading of printed text.  We do not know what significance this fact may have, relative to other modes of speaking.

### 5.3. Recording environment

The acoustic environment in the XRMB laboratory was far from ideal for speech recording.  The room itself was rectangularly shaped, approximately 12.2 x 5.8 x 3.7 m.  Two walls were flat and hard, and two were covered extensively but irregularly by large fabric-covered acoustic absorption panels.  The floor was thinly carpeted, and the ceiling surface irregular in shape and texture.  A significant portion of the room was occupied by the lead-shielded beam line, scintillation counter, workstations, and instrumentation racks.  The speaker's chair was surrounded on three sides and above by 1.2 x 1.2 m acoustic shields.

The instrumentation scattered about the lab, and the HVAC ducts under the floor, supplying the laboratory itself and an adjacent computer room, generated a fairly noisy background SPL of about 60 dB (C-scale).  Speakers who weren't shy produced signal-to-noise ratios for the acoustic track of around 30 dB.  Shy speakers (quite a few) produced poorer ratios.  Given the time and resources available to this project, these could not be improved.

# CHAPTER SIX
## Post-processing and Evaluation of Pellet Position Histories

Raw pellet-position histories from each speaker were inspected and treated off-line in several ways after their initial acquisition. Post-processing and evaluation procedures were intended to maximize accuracy and reliability of these signal streams, and to enhance their accessibility by standardizing representational and file-storage conventions.

### 6.1. Target and image-plane corrections

The rectangular coordinates assigned to an evaluated pellet center during real-time tracking are proportional to electron beam deflections that generate the X-ray beam whose path the pellet interrupts. Those deflections correspond in turn to two-dimensional coordinates on the surface of the XRMB system's tungsten target, where the X-ray beam originates. The target itself is not flat, but cylindrically shaped, flat from side to side (in the system's x-direction), but gently curved from top to bottom (in the system's y-direction), with a radius of curvature of 0.5 m. A numeric correction for distortion associated with this curve was the first step in post-processing raw pellet-position data. The correction amounted to sample-by-sample x and y-direction gains, directly proportional to the observed y-coordinate of each pellet. Signal and measurement-error components of each observed position are affected by this correction in the same way.

The XRMB operates something like a pinhole camera, after a fashion illustrated in Figure 6.1. The beam of X-rays generated at a point of intersection of the electron beam and system target, flattened in this view, exits the system through a pinhole positioned a known distance from the target. The beam follows its course toward pellets that are assumed to lie in an image plane at some specified distance from the pinhole. The coordinates assigned to a pellet represent the image-plane projection of its position, according to the principle of central projection. The ratio of each assigned coordinate value in the image plane to the corresponding known coordinate on the target is equal to the ratio of image-plane-to-pinhole and pinhole-to-target distances. The pinhole-to-target-center distance is physically fixed. In contrast, the image-plane-to-pinhole distance must be measured during each recording session, and may vary from about 300 to 800 mm, allowing for some flexibility in locating the speaker within the system image field. The image-plane-to-pinhole distance measured for each speaker was used to numerically scale pellet position data to life size.

### 6.2. Equal-time interval conversion and re-sampling

Each raw pellet-position history contained a series of two-dimensional coordinates, representing found centers to the nearest possible grid location (see Section 7.1), and an accompanying series of explicit time stamps defining, to the nearest microsecond, the moments when pellet centers were found. Pellet coordinates and time stamps were generated by a recognition computer, operating in parallel with other processors whose general function was governed by real-time sampling-control code.

The sampling-control software governing pellet tracking introduced three inconvenient temporal features into raw pellet-position histories.

Figure 6.1: Geometric principles of XRMB function
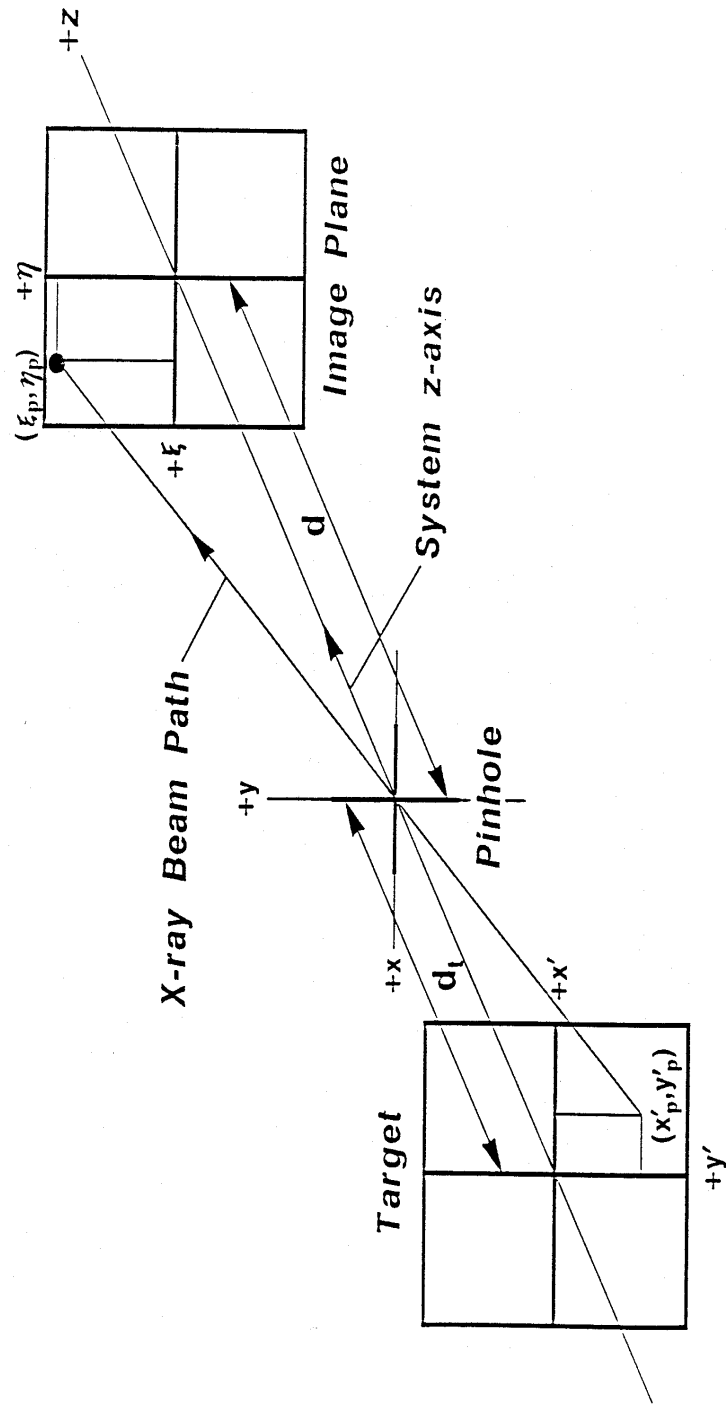
1) Time-stamps associated with sampled positions of any specific pellet were different from the time stamps associated with other pellets. That is, the position of any pellet was known at a unique time, different from the times at which the positions of any and all other pellets were known. This sampling-schedule effect is logically necessary. When multiple channels are sampled sequentially, only one channel can be sampled at any specific time. But, the effect is inconvenient for certain post-processes that might be applied to the data. For example, a coordinate-system transformation of the kind described in Section 6.3., used to re-express articulatory pellet positions with respect to cranially defined axes, requires that all pellet positions be sampled at the same rate, and known at the same times.

2) The sampling control code introduced a modest, systematic sample-to-sample jitter in the inter-sample intervals of the most frequently sampled pellets. This feature of the data presents problems for other post-processes like digital filtering, which are well-defined only for equal-time-interval signal streams.

3) The interval from time zero within each record to the first recorded sample for any given pellet, was greater than zero for all pellets. During this time-to-first-sample interval, pellet positions could not be known.

The first two features of the raw pellet-position histories were addressed by interpolation and re-sampling. Target and image-plane-corrected data streams were fit with piece-wise continuous smoothing splines, then re-sampled at a uniform rate of 160 samples/second, at common times, beginning with zero time within each record. For all pellets except T1, this process increased the effective sample rate above the original (nominal) tracking rate. For T2-T4 and LL, the nominal rate was increased by a factor of two, from 80 to 160 samples/s; for other pellets, the rate was increased by a factor of four (or eight), from 40 (or 20) to 160 samples/s. The re-sampling procedure assigned out-of-range coordinate values to pellets during their respective zero-time-to-first-sample intervals, effectively removing any requirement for a time-to-first-sample parameter associated with each pellet-position history in each record. All other newly computed coordinates, for every pellet in every record, were based on the observed coordinates obtained during real-time tracking.

The derived equal-time-interval, same-sample-time format should prove convenient for future users of database materials. This format will circumvent potential problems associated with displaying differentially sampled data, and with processing operations that require equal-time/same-time samples.

*6.3. Coordinate-system transformation (from machine to head-space)*

The coordinates assigned to pellets during tracking represent pellet positions in the system image plane, and are therefore defined with respect to measurement axes intrinsic to the XRMB instrument (i.e., defined with respect to "machine space"). We prefer, however, to define positions of the mobile articulators (e.g., the tongue, mandible, and lips) relative to reference axes fixed with respect to the speaker's head, even though the positional measurements of both the articulators and the head were originally made relative to axes fixed external to the speaker.

Figure 6.2: Schematic illustration of coordinate-system transformations
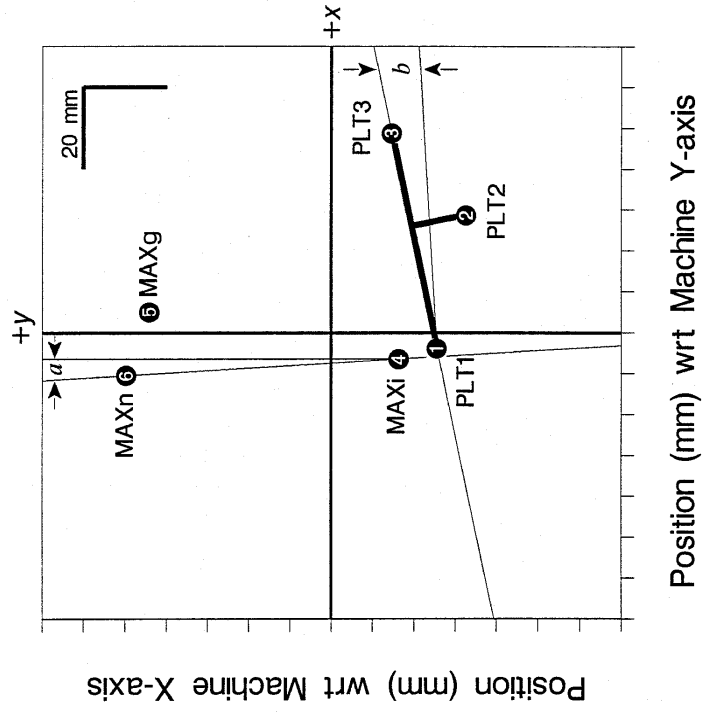
A two-stage transformation, illustrated in Figure 6.2, re-expressed position histories of articulator pellets relative to a two-dimensional, anatomically-defined cranial coordinate system appropriate for each speaker's midsagittal plane. The first stage involved translation (e.g., redefining pellet 4 in Figure 6.2 as the coordinate system origin) and rotation (by the angle $a$) of coordinates of each pellet, sample by sample within each interpolated and re-sampled record, to an intermediate origin defined at one fiducial pellet (usually MAXi), and a reference axis defined by that and one other fiducial pellet (usually MAXn, but sometimes MAXg). This transformation removed rotational and translational components of head motion, reflected onto the system image plane, from the trajectories of all pellets, effectively re-expressing their positions relative to an arbitrary coordinate system fixed with respect to the speaker's head. The second transformation involved an additional translation (making pellet 1 the origin)and rotation (by angle $b$). These parameters were derived from each speaker's biteplate tracking record (see Section 5.2.2.1.), and represented the fixed geometric relationship between fiducial and biteplate pellets. This second transformation established the coordinate system origin at the caudal-most edge of the central maxillary incisors, an x-axis corresponding to the intersection of the midsagittal and maxillary occlusal planes, and a y- axis normal to the maxillary occlusal plane and passing through the origin. By convention, in this anatomic coordinate system, the +x and +y directions are unambiguously defined to be out of the mouth, and toward the top of the head, respectively.

A standard coordinate system for describing speech movement, established in the same way for each speaker, is one of the underlying strengths of the speech production database. A common representational convention of this type facilitates meaningful comparison of data from one speaker to the next, by minimizing spatial and temporal biases that can result from variably placed coordinate-system axes ( Westbury ,1994).

### 6.4. Head-position evaluation

Head position is an important variable in XRMB experiments, in large part because the head is left free to move while speakers talk and perform. Certain head movements will introduce additional measurement error to the acquired signal streams, because they violate assumptions inherent to the system's operation. Measuring head position provides a method for limiting this error, and for estimating confidence limits that can be associated with the acquired data. Two types of off-line measurements of head position were made for each speaker.

### 6.4.1. Three-dimensional estimate of head position from calibration records

A quantitative estimate of position and orientation of each speaker's midsagittal plane, relative to a three-dimensional machine-based coordinate system, was based upon data obtained from the biteplate tracking record (see Section 5.2.2.1.). Mean image-plane-projected coordinates of biteplate pellets were used together with known distances between them to estimate three-space coordinates for each pellet. These data were used in turn to determine the position of the speaker's midsagittal plane relative to the specified image plane. They were also used to determine Euler angles relating orientation of the midsagittal plane to the XRMB system image plane. The median magnitude of off-plane head position, across all datasets from all speakers, was approximately 6 mm, and ranged between 0.1-22 mm (corresponding to a median

60

magnitude scale error of slightly more than 1%, and range of 0-4%). Median magnitude yaw and roll angles (about rostral-caudal and dorsal-ventral axes, respectively), relating orientation of the local (three-dimensional head-space) coordinate system to the global (three-dimensional machine-space) coordinate system, were 2.6 and 1.8 degrees, respectively. Off-normal orientations of this magnitude introduce scale errors of less than 1%.

## 6.4.2. Scale effects of head position, estimated from fiducial inter-pellet distances

Data from biteplate tracking records were used to approximate true distances between fiducial pellets. These distances were compared to observed distances, sample by sample within every record, to qualitatively evaluate the effects of changes in head position during task performance. Record histories of average inter-(fiducial)-pellet distances, and derived scale errors, were routinely generated to judge constancy of head position across records. Examples of histories of this type, for two different speakers, are shown in Figure 6.3. Speaker JW34, shown in the lower half of the figure, remained relatively still over the course of her recording session, with her head positioned about 6 mm further away from the system image plane than specified, resulting in a scale error of slightly more than -1%. Data from speaker JW16, shown in the upper part of the figure, indicate two abrupt changes in head position, involving roughly an 8-mm shift between records 59 and 60, and a 20-mm shift between records 93 and 94. The latter shift corresponded to a scale error change of almost 4% between adjacent records, such that the pellet data appeared almost 1.5% greater than true prior to the shift, and 2.5% smaller than true after the shift.

Measures of relative head position, sample to sample and record to record, were too imprecise to allow full three-dimensional compensation. Consequently, scale error histories were used only in a simplified way to compensate numerically for varying head position, within and across records: Scale changes associated with head motion were treated as simple translations along the XRMB-system z axis; these were compensated by scalar multiplication of each pellet-position sample within each record by the corresponding sample of the reciprocal of the ratio of observed-to-true interpellet distance for the most widely spaced fiducial pellet pair.

## 6.5. Synchrony evaluation

Speech production involves multiple actions, at many physical levels. Understanding the control and intent of this rich motor process depends on our ability to concurrently record, view, and analyze as many relevant signal streams as practically possible. The XRMB system was designed to meet this need by rapidly recording the time-varying positions of articulators in synchrony with other physiological data. The reliability of inferences drawn from these data will depend significantly upon the accuracy of inter-channel synchronization.

An interrupt flaw in the XRMB acquisition hardware, randomly affecting 5-10% of records obtained from each speaker, was discovered after all database speakers had been recorded. This flaw had the effect of introducing a time lag of both physiological channels, relative to all pellet-position channels, equal to the duration of one packet (2048) of samples times the aggregate physiological sampling period. Fortunately, the cue-tone latency recorded at the start of each record provided a reliable marker for this recording error. Correct alignment

was restored by advancing in time all physiological tracks relative to all pellet tracks, by an appropriate interval.  The results of this procedure were subsequently verified by

Figure 6.3: Scale error histories (across records) derived from inter-(fiducial)-pellet distances, for two speakers
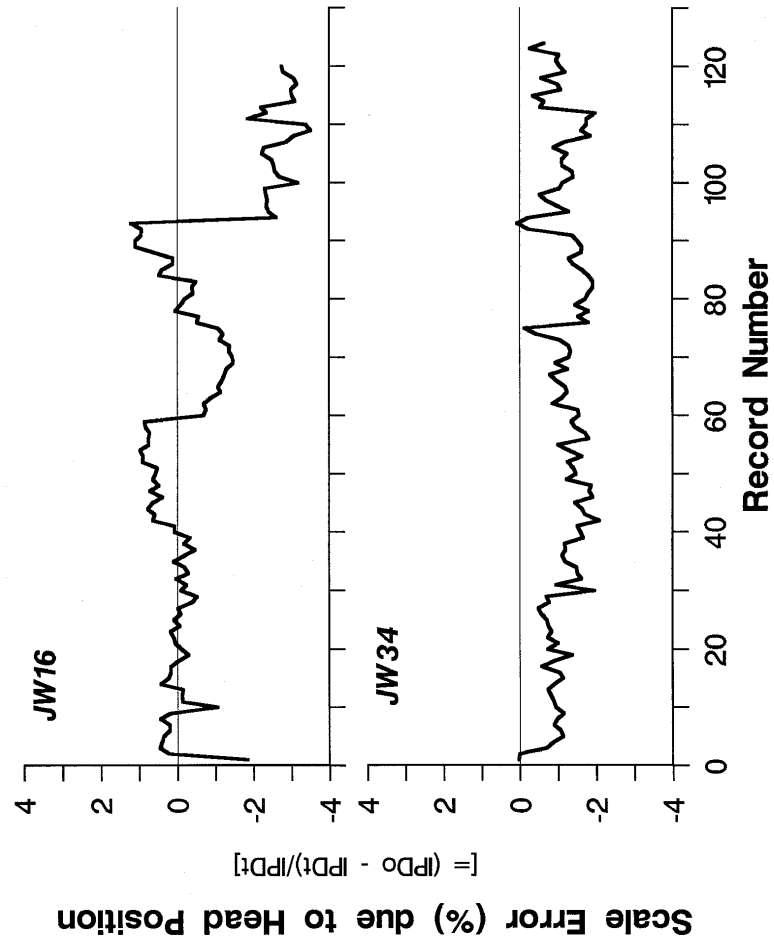
calibration experiments designed to evaluate temporal resolution of the XRMB system (see Section 7.2). A list of records by speakers repaired according to this procedure is contained in a short text file accompanying the database.

A very small proportion of records (approximately 0.5%) were lost to a more destructive synchronization error that we do not yet understand. These records have been declared invalid and stripped from the datasets that originally contained them. They have also been listed in the text file identifying modified records.

Records from datasets other than our own, recorded at the XRMB facility over its several years of operation, may have been affected by synchrony errors, though we have little knowledge of this. The only public comment on the problem that we have seen appeared in a recent article by Papcun et al. (1992), in reference to a single record.

*6.6. Alignment of palate outline and pellet trajectories*

There should be good agreement between palate outlines (see Section 5.2.2.4.1) and certain extreme positions of tongue pellet trajectories above the MaxOP, since some parts of the tongue routinely contact the palate in certain circumstances. Sagittal-plane pellet trajectories shown in Figure 5.6., from a sentence-type record from one speaker, show reasonably good agreement with respect to the palatal outline. In database records, tongue-to-palate contact is unequivocal whenever lingual fleshpoint trajectories intersect the palatal curve. (Intersection is possible because the centers of pellets used to define the palatal outline, and the positions of lingual fleshpoints, are respectively one-half a pellet diameter below and above the surfaces they represent.) There may also be tongue-to-palate contact when pellet trajectories fall short of the palate, but by portions of the tongue other than the fleshpoints that are tracked. However, pellet trajectories should not penetrate the palatal curve.

Good agreement between lingual pellet trajectories and palatal outlines did not always occur. Instances of poor agreement tended to be of three general types. The first of these is apparent in Figure 5.6., where dorsal positions of the T1 pellet seem to fall short of the palatal curve, even though there was good reason to have expected tongue-to-palate contact. These instances may well involve contact, but not at the pellet's observed position.

A second type of poor agreement, common for trajectories of dorsal lingual pellets T3 and T4, involved periods when the pellets appeared to penetrate the palate outline, particularly for outlines generated from the stone model of the maxillary arch and palatal vault. There is a simple, plausible explanation for this type of disagreement. Palatal outlines generated from stone models likely represent a lowered position of the soft palate. The procedure used to obtain dental impressions requires that subjects breathe nasally while holding a mouthful of alginate. Trajectory penetration of the dorsal aspect of palate outlines was less common for outlines derived from *in vivo* traces, where subjects presumably traced the palate with the velum in a raised position while breathing orally. At this time, we consider a modest amount of dorsal penetration of model-generated palatal curves to reasonable and acceptable.

A third type of poor agreement between pellet trajectories and palate outlines involved minor, but perceptible penetration by ventral tongue pellets. We believe this results from coarse spatial sampling inherent to the (stone) model-based procedure for defining the palatal curve.

64

Locations of pellets placed along the model palate midline were clearly identifiable from the scan image (e.g., see Figure 5.7.), but the shape of the palate between pellet locations is not captured well. The scan-generated approximation of the palatal curve is piece-wise linear, and necessarily somewhat inaccurate in regions where the curve changes rapidly (e.g., in the vicinity of the alveolar ridge).

For those speakers with manual *in vivo* palate traces **and** model palate outlines, pellet trajectory extrema tended to match the manual trace better, but not consistently so. For example, the model palate outline for JW59 fit the tongue pellet trajectories better than the hand trace. This was likely the result of poor tracing by the subject or experimenter. Thus, it should not be assumed that one method of palate outline generation is necessarily superior to the other. There are disadvantages associated with both methods. Manual traces are often limited in their dorsal excursion (depending on the queasiness of the subject), and may follow inconsistent paths along the palate. Model-generated traces sometimes fail to fit the tongue trajectory data well, though it is easier to deduce the reason than in cases where a manual trace produces a poor fit.

## 6.7. Mistrack identification and annotation

Not all pellets were properly tracked all of the time by the XRMB system. Instead, some pellets were momentarily lost during some tracking intervals. At those times, pellets were said to have been *mistracked*, and the pellets' assigned position coordinates will be wrong. Correct data are missing during such intervals, and cannot be recovered.

Typically, pellets were not mistracked for the full duration of a record, though sometimes this happened. It was more common for them to become lost only for a short while (ca. 50-500 ms), often near the beginning of a record, and then to be recaptured at some later point. For this reason, only the affected portions of pellet-position histories need be noted and ignored. An exception to this rule involves mistracking of fiducial (and mandibular) pellets, when these might subsequently be used to define coordinate systems within which other "articulatory" pellets are said to move. Coordinate system re-expressions are impossible if critical pellets are mistracked, and it then becomes necessary to invalidate all data obtained during the affected interval.

The proportion of data lost to mistracking, expressed relative to each speaker's total tracking time multiplied by the number of pellets tracked, was low, averaging only 1.9% across a subsample of eight speakers (JW7,12-16,18-19). The number of *records* (in contrast to the amount of *time*) involving some mistracking, on one or more pellets, was proportionally higher, particularly for some speakers. Probably no more than 65% of records were obtained without any evidence of mistracking on at least one channel. However, mistracking was rarely a justification for repeating records during acquisition. Insisting that all tracks be recorded perfectly cleanly would have been prohibitively expensive, in terms of risk and time.

## 6.7.1. Probable causes of mistracking

A simple description of XRMB system operation provides some understanding of how and why mistracking occurs. During successful tracking, each pellet is followed by a labeled flying-spot X-ray scan, 6 mm on a side. The moment-to-moment location of this scan, referred

to as a *raster*, is under computer control, and depends on the pellet's earlier found positions.  At discrete moments in time during an ongoing record, the position of the pellet is determined when it falls within the area covered by its associated raster scan, and casts a recognizable shadow on a two-dimensional X-ray count registered by a NaI crystal detector.  The coordinates assigned to each sample of the pellet's position represent the centroid of its shadow.

If no shadow, or an unrecognizable shadow, falls within a raster scan, the XRMB system returns a *Not Found* (NF) error message, and increments an accumulating mistrack variable associated with each pellet label.  The total number of NF samples per pellet, per record, is automatically logged on line as a dataset is acquired.  The system also responds in two additional ways: first, by re-directing the relevant raster scan to return to the area surrounding the pellet's initial found position in the record, in the hope that it will pass nearby and be captured again; and secondly, by returning false coordinate values for the pellet, corresponding either to the coordinates of its first-found position in the same record (for all sets acquired before JW49), or to fixed coordinates well outside the system image field (for all later sets).  Presumably, the system fails to find a pellet during one or more successive samples of its position either because the change in pellet velocity between previous and current samples is so great that no shadow appears within the current raster; or, because X-ray absorption by adjacent structures causes an insufficient contrast between the signal shadow cast by the gold pellet, and background shadows cast by tissue, bone, teeth, and fillings.

A pellet may also be lost to the system, and assigned incorrect position coordinates, when the labeled raster scan initially assigned to it begins to follow another of the pellets placed within the image field.  This can happen when the trajectories of two pellets come close enough together so that their positions project shadows in two (or more) raster scans closely adjacent in time. If a wrong shadow is then selected for evaluation, the position coordinates assigned to the raster's pellet label will also be wrong.  The usual outcome in this situation, typically referred to as *raster hopping*, is that two raster scans begin to follow one pellet, and subsequent positions of the other lost pellet are no longer captured.  The position time histories associated with both pellet labels will then be essentially the same, at least until some subsequent moment when both scans again overlap an area including both pellet shadows.  At that point, a newly correct re-association of labeled raster scans and pellets will then become possible, and may in fact re-occur.

A final way that a pellet can momentarily be lost to the system is related to the phenomenon that causes raster-hopping between pellets, but involves false recognition of one or more (complex?) shadows that might appear within a raster scan.  When the labeled scan associated with a pellet encounters a shadow that is sufficiently pellet-like (i.e., with sufficient contrast between radio-dense and light areas), but cast instead by some other relatively dense object (e.g., dental filling or unusual tissue density), the system may falsely conclude that (only) a pellet has been found, and wrongly return the coordinates of the shadow centroid.  Of course, these coordinates will then be in error, though it is also possible that they will initially appear to be quite plausible, since they first arise from a shadow that is close to the pellet's true position.

False recognition may form the basis for tracking problems that plagued the beginnings of many database records.  The system often seemed to initialize records, indicating that all initial pellet positions were known, even though it quickly became apparent after initialization that

successful tracking had broken down.  In cases of this type, which were referred to as *false initializations*, one or more of the labeled raster scans remained relatively stationary after the record had begun, presumably because the system was falsely tracking a stationary shadow not cast by a pellet, or otherwise locked onto a false initial pellet position.  False recognition may have been related to tracking difficulties associated with pellets that moved about in relatively dense local environments.  It was frequently true among database speakers that pellet tracking in the vicinity of the tongue dorsum (T4) failed more frequently than at other oral locations.  Data for T4 pellets were characteristically noisy, and lost raster scans were common, even when the pellets themselves seemed to be moving relatively slowly and along simple paths.  It is plausible that these problems resulted from a form of false recognition, given that the X-ray beam aimed at these pellets had to pass through thicker layers of tissue, resulting in greater absorption and more diffuse shadows whose centers were less clear.  Unfortunately, there are no solid criteria that can differentiate falsely recognized and evaluated pellet shadows, and falsely specified pellet positions, from those that are correctly recognized, evaluated, and specified. Accordingly, some part of every X-ray microbeam dataset may be indecipherably wrong.  We have to believe it is an insignificant part.

*6.7.2. Locating and treating mistracks*

Intervals where pellets were mistracked were interactively located and annotated, for each pellet, record, and speaker.  Mistracks associated with raster-hopping were located by visually scanning successive positions of problematic pellet pairs, or pellet sets, linked by line segments. Mistracks representing not-found conditions were located by identifying intervals where pellets remained too long at first-found and/or extreme, out-of-field positions.  Figure 6.4 illustrates an example of  a mistrack identification in a single pellet position time history.  A sample mistrack file for one speaker is shown in Appendix E.

The lengthy process of identifying and marking mistrack intervals required something on the order of two person-years, distributed among members of the project team.  In fully post-processed signal streams, out-of-range coordinate values have been assigned to each pellet, during each record, during all intervals judged to have been mistracked.  Simple thresholds can therefore be used to distinguish well-tracked and mistracked data, in those waveforms.

Figure 6.4: An example of a mistrack
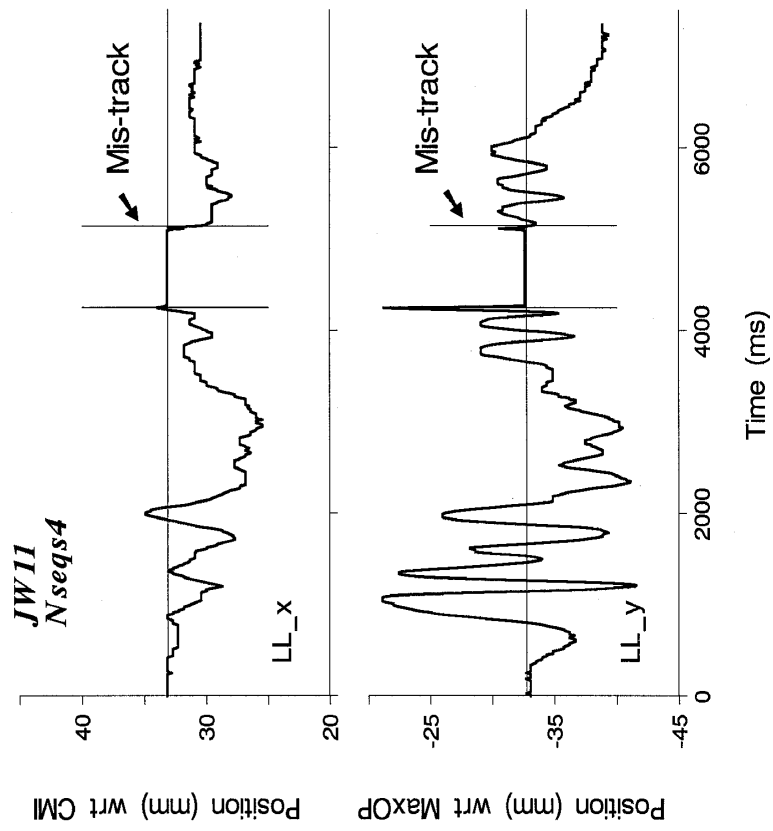
# CHAPTER SEVEN
## XRMB System Performance Characteristics

The XRMB system is fundamentally a measurement device which we use to determine the relative spatial positions of radiodense markers lying within its image field.  In simple terms, what we want to learn from such a system is where articulator pellets are, and when they are there, as speakers move to speak.  Thus, it is important to know how well the system measures position and time.  These performance features determine in large part what legitimate interpretations we may make of any data we obtain.

It is also important to understand that the system performs its function by exposing speakers to ionizing radiation.  There is risk associated with the method, though it is difficult to quantify precisely.  Conventional descriptions of this risk, of the type provided to potential speakers who must make informed decisions regarding participation in experiments, are expressed in terms of entrance dosage.  A brief summary of dosage estimates is described below in Section 7.3.

### 7.1. XRMB spatial resolution

The XRMB locates spherical pellets within its field by "pointing" an x-ray beam at them, and then assigning positions according to the location of the point of maximum absorption.  Tracking pellets requires that the X-ray beam be stepped in time along the system target, and thus, across the system field.  In the limit, the electron beam of the XRMB can be stepped along the system target in integer DAC (digital-to-analog controller) steps that correspond to 0.0625 mm along a flat surface 60 cm distant from the electron-beam deflection coils.  In principle, this beam can therefore be positioned at spatial intervals no smaller than 1/16th of a millimeter along the target surface.  By definition, pellets can be found only at locations where the beam can be pointed.  Thus, under optimal conditions, for an image plane 60 cm distant from the system pinhole, pellet centers at every sample can be found, and distances between them resolved, at spatial intervals separated by no less that 0.0625 mm.

Spatial resolution in this system is inversely proportional to the distance of the image plane from the system pinhole.  For database speakers, that distance was typically on the order of 53 cm.  Consequently, for their sessions, we can estimate system resolution to have been approximately 13% better than that allowed by whatever DAC-step size might have been applied during pellet tracking.  In the tracking mode used for all database sessions, the minimum step deflection for the electron beam was restricted to eight (8) DAC steps. This restriction meant that from sample to sample, each pellet center could therefore be found only at one of many "grid" locations in an image plane partitioned at approximately 0.44 mm intervals, in both x and y-directions.  In Figure 7.0., two hypothetical pellet locations are illustrated, within a 5.28 x 5.28 mm raster.  The pellet with the heavy outline, centered on a grid location, will most probably be found only at that location, for repeated samples of its stationary position.  The pellet with the lighter outline, centered off of any possible grid location (at the position indicated by the unfilled circle), can be found at any of four grid intersections surrounding the true location of its center, for

Figure 7.0: Effect of relative locations of pellets and the the system grid for pellet center assignments
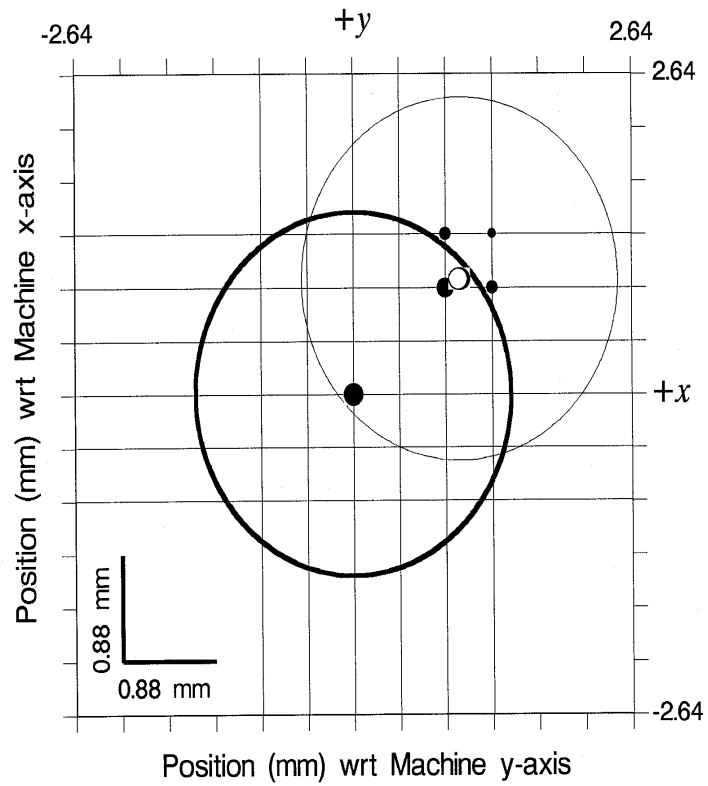
repeated samples of its stationary position. The grid location closest to its true center location will be returned most often.

From calibration experiments performed on the system, we know that repeated measures of the positions of stationary pellets show their perceived centers to occur only at grid locations, but not the same locations, from sample to sample. Rather, any specific pellet's found center will typically be distributed across at least two possible positions in both x and y directions (cf., Figure 7.1), presumably because there is some temporal variation in (a) photon density per image-field unit area generated by the XRMB system, and/or (b) energy absorption per unit volume distributed across the pellet mass being tracked.

We also know that the shape of the distribution of repeated position samples for a stationary pellet depends upon its proximity to image field gridlines. A pellet whose center falls truly midway between gridlines is found about half of the time at either adjacent grid location; a pellet whose center is truly on a gridline is rarely found elsewhere. Thus, from sample to sample, a pellet may be found where it isn't. In fact, this must be true for every sample for a pellet not on a gridline. But, even when not on a gridline, the pellet's true position will influence the probability that the pellet will be assigned a position at either neighboring gridline. Thus, the positional accuracy of the XRMB system, measured in terms of the discrepancy between observed and true positions of pellet centers, varies according to the relative locations of measurement gridlines imposed by system operation, and the pellets that traverse them. A simple, broad statement of measurement error that generalizes easily across all conditions is difficult to make or interpret. Calibration measures across a series of some 40-odd static positions, each separated by 0.05 mm intervals, and spanning five adjacent image-field gridlines, have shown that RMS positional error, per sample during stationary tracking, averages approximately 0.15 mm, and ranges between 0.030-0.250 mm, when electron-beam deflection is constrained to eight-DAC-step intervals.

Positional error during non-stationary tracking has never been documented, though there is reason to expect the error to be somewhat greater, and perhaps even velocity-dependent. The position of a moving pellet will change during the time required to generate its tracking raster (1.44 ms, for a 12x12-pixel raster, 6 mm on a side). A pellet moving 400 mm/s in one dimension (on the high side for speech, but not impossibly high) will traverse almost 0.6 mm, across one grid interval, merely in the time required to produce the raster. This motion should stretch the shadow cast by the pellet, along the direction of motion, and might well affect the system's perception of its position.

*7.2. XRMB Temporal Resolution*

Sampling functions involving physiological channels and pellet-position histories are governed by a clock that is accurate to 0.5 microseconds. Explicit time stamps, to the nearest whole microsecond, are associated with raw samples of all pellet positions. Time stamps must be *explicitly* associated with each pellet sample because samples are not all taken at equal time intervals, nor at the same times (see Section 6.2), during data acquisition. The time stamps are subsequently discarded by post-processing operations

Figure 7.1: Distribution of found centers: conventional tracking of a stationary pellet
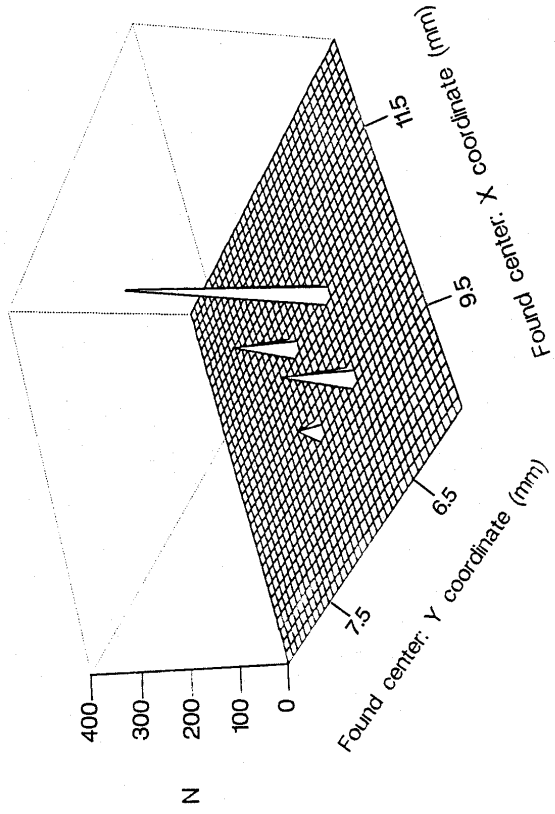
which approximate raw histories by *derived* waveforms that are equal-time-interval and "simultaneously" sampled.

A coarse empirical test of cross-channel synchronization between pellet and physiological data, involving fast-tracking of pellets (@ 250 samples/s) attached to a tapping wand used to strike the wind-shield of a dynamic microphone, confirmed that the two data types were correctly synchronized to within 2 ms (one-half the highest-pellet-rate sample period). This estimate of synchrony represented the lower limit permitted by the method; presumably, better performance would have been confirmed for faster pellet tracking.

This same test was also conducted to develop an understanding of, and remedy for, a cross-channel synchrony flaw introduced by acquisition hardware during some records recorded from some database speakers (see Section 6.5.). During this test, samples of properly-synchronized and flawed records were collected, using methods analogous to those employed during database recording sessions. Data of both types were subsequently trigger-averaged to establish three important facts bearing on cross-channel synchronization within records:

(1) Correct cross-channel synchronization is reliably indicated by a record-level feature referred to as a **short** *cue-tone latency*: an interval between the first sample of the sound pressure wave, and onset of the speaker's starting cue that may be seen at the beginning of each acoustic track in each record. Short latencies range between 8-23 ms, and their specific durations vary randomly from record to record. The specific values of latencies falling within this range have no effect whatsoever on physiological and pellet-track synchronization. All latencies falling within this range indicate cross-channel synchronization that is correct to within the measurement limit of the calibration test.

(2) Moderate cue-tone latencies, in test and database records generally greater than 55 ms and less than 70 ms, indicate synchrony flaws whose effects are also independent of variation in the specific value of the latency.

(3) The effects of synchrony flaws of the type described in (2), above, can be completely eliminated, to within the 2-ms error window, by a phase-shifting technique which advances all physiological channels by an amount of time equal to the product of 2047 and the sampling period corresponding to the aggregate physiological-channel-sampling rate.

For exploratory purposes, cue-tone latencies were measured for a sample of 1122 raw database records, distributed across portions of data from 13 speakers. In 1035 (92%) of these records, latencies were of the short, acceptable type, with a mean value of 14.5 ms, and minimum and maximum values of 8.15 and 23.89 ms, respectively. In these records, cross-channel synchronization can be assumed to have been originally correct. In 82 records (7%), latencies ranged between 55.9 and 66.2 ms. These were repaired using the procedure described in (3), above, and have been retained in the database, now presumably properly synchronized. A small number of records (5 of 1122) revealed very long latencies, of an unknown type, and have purposely been deleted. No cue-tone-latency types other than these were observed.

## 7.3. Dosimetry Estimates

The XRMB technique uses a narrow beam of x-rays, roughly 0.4 mm in diameter, to track pellets in real time.  The technique therefore exposes speakers to ionizing radiation.  Factors that limit this exposure in significant ways include: (1) small raster exposures only of tissue immediately adjacent to expected locations of pellets; (2) sampling rates no more frequent than necessary for reliable tracking; and, (3) a short (10 microsecond) dwell-time of the x-ray beam per unit area.  Tracking a pellet at a rate of 100 samples per second, for 1 minute, exposes tissue in the 30-40 mm$^2$ area surrounding the pellet center to the x-ray beam for 60 ms.

Aggregate entrance dosage is affected by the energy of the electron beam, the level of emission current, time of exposure, and pellet sampling rate.  Dosage calculations for the UW XRMB system, confirmed to within 10% by direct measures using thermoluminescent detectors and scaled to whole body dose equivalent (derived from NCRP Report #91, Table 5.10), indicate maximum integrated dosages of 97mR for a total tracking time of 1200 seconds with an energy of 450 kV at 1.75 mA and a sampling rate of 160 samples/s.  This integrated entrance dosage is comparable with lower-energy diagnostic radiological procedures (~70 kV) such as dental bitewings, or upper thigh X-rays, that speakers may have received from other sources.

Potential speakers were routinely encouraged to decline to participate, or to discontinue participation, if they had any misgivings about risks inherent to the procedure (cf. section 4.1). Some did so; the majority did not.  Speakers were paid a nominal amount for their participation. All procedures involving human subject participation in the development of the XRMB Speech Production Database were reviewed and approved by an Institutional Review Board of the University of Wisconsin.

# CHAPTER EIGHT
## Dataset Organization and Contents

The following information is necessarily preliminary in nature. This handbook is intended for distribution with subsets of material from the XRMB Speech Production Database. At the current time (June, 1994), no platform-transparent version of the database is ready for open release. For that reason, many important details about database organization and format are not yet known. This brief description will be revised and expanded, as necessary, to be consistent with future releases of database materials.

Broadly, we can guess even now that the XRMB Speech Production Database will be divided into information units of different sizes and types. The largest of the primary units will be the speaker dataset. Each dataset will be divided in turn into a series of records, with each record separating into a collection of synchronous time series. Supplementary collections of information, chiefly in the form of text files and simple space-delimited tables, will also be included. Some (but not all) of these will span speakers, and should typically be limited to convey one type of data.

The most plausible target medium for database distribution will be compact disk. CD hardware and control software are common and relatively inexpensive, and provide quick random access times for relatively large bodies of data. The disks themselves are sturdy and high-capacity. The total database requires something on the order of 6.0 Gbytes of mass storage, with perhaps 95% of that total required for the physiological (wide-band) channels.

## 8.1. Speaker Datasets

All time series data recorded from each speaker are contained within a dataset, corresponding to a subdirectory identified by a character string of the form JWNNN (where the last three characters are filled by speaker ID numbers ranging from 5-63, and including three 3-digit numbers 382, 472, and 502). There are sixty speaker datasets, but only fifty-seven different speakers. Three women in the sample are represented twice, by different datasets recorded on different dates, with at least one of their respective datasets significantly incomplete. Thus, for these three speakers (JW17/22, JW38/382, and JW47/472) there is some duplicate material. Duplicate recording sessions were motivated by XRMB system malfunctions that prevented full sessions on each speaker's first visit. Follow-up sessions were run 17, 8, and 42 days, respectively, after the first session for these speakers.

## 8.1.1. Records

Datasets are divided into a series of **record** subdirectories, each containing the set of time-series channels recorded during a single continuous task interval. In the ideal case, there should be 118 (or 119, for some subjects from whom data were obtained near the end of the acquisition period) records within a dataset, representing one recorded example of each task within the database task inventory. Generally, however, there are different numbers of records within each dataset. For some speakers, some records were recorded more than once, in response to mistracking or other acquisition flaws. For other speakers, records were damaged and lost

during acquisition, due to system errors. Other records may have been deleted during post-processing and evaluation procedures because they were judged to be unacceptably flawed or uninterpretable. However, no records were deleted from any set due to speaking or performance errors. All data were obtained at some risk to speakers, and are treasured accordingly. (A master list of surviving records, by task across speakers, is provided as an appendix. Lists of surviving records, by speaker, will be stored in a subdirectory of supplementary files.) The names of records indicate to some extent record content according to task type. All records containing citation words, for example, have names beginning with the character string *words*.

*8.1.2. Channels*

Each record subdirectory contains a collection of files representing the time series data obtained during the record. In the typical case, 19 files will be found: one each for the sound pressure wave and neck wall vibration; one each for the x and y coordinates of each of eight articulator pellets; and one representing a vector of explicit time stamps appropriate for any of the sixteen pellet-coordinate histories. The names of channels indicate channel content. For example, *tacc* is used to indicate the neck-wall-vibration signal stream; *t1y* the y-coordinate of the pellet positioned in the vicinity of the tongue blade; and *mix* the x-coordinate of the pellet glued to the mandibular incisors. All channels are equal-time interval, with sampling rates of 21739 (SPW), 5434 (NWV), and 160 (pellet) samples per second. (Sample rates for SPW and NWV channels are different from these values for three speakers: JW7-9. Specifications for those speakers are available in Chapter 5 of this Handbook.) The time spans of different channels are typically different, though the first sample of each channel can be considered to start at zero time within the record. For some speakers, fewer than 16 pellet-coordinate channels are represented, because fewer than eight pellets were tracked during the record. Pellets sometimes came loose during a recording session, and were removed; at other times, pellets were purposely removed because their trajectories overlapped those of adjacent pellets, causing unacceptably high levels of mistracking.

The position time histories for fiducial pellets attached to the nosebridge, maxillary incisors, and other relevant reference locations, have not been retained in the channel collection of each record. These channels were used to re-express the locations of articulator pellets relative to an anatomically standardized coordinate system appropriate for each speaker's head. Their original locations were of course fixed relative to the speaker's head, and thus, their transformed positions do not move.

Pellet-coordinate channels are ASCII, in calibrated millimeters rostral/superior (positive) or caudal/inferior (negative) to the maxillary occlusal plane (for y coordinates), and ventral/anterior (positive) or dorsal/posterior (negative) to the central maxillary incisors (for x coordinates). Sound pressure wave and neck wall vibration channels are binary in arbitrary amplitude units.

*8.2. Supplementary Files*

A number of subdirectories, lateral to datasets, hold useful information about database speakers and signal streams. Generally, the names of these subdirectories will suggest their content. Where appropriate (e.g., for vocal tract outlines), each subdirectory will contain a group of files, named in part according to speaker identification number. Other subdirectories contain only one or a few files that generalize across all speakers (e.g., the ASCII DARPABET transcription of the task inventory, and derived phone diads).

*8.2.1. Vocal Tract Boundary Outlines*

Files holding palatal curves and posterior pharyngeal wall outlines are contained in a single subdirectory. There are two ASCII files per speaker, named according to speaker identification numbers, one for the palate (.pal) and one for the pharyngeal wall (.pha). The former is represented by space-delimited vectors of x-coordinate locations, and y-coordinate locations, along the palatal surface. The latter is represented in a similar fashion, but by only two ordered (x,y) pairs, corresponding to a straight line approximation of position and orientation of the posterior pharyngeal wall, spanning some distance on either side of the maxillary occlusal plane.

*8.2.2. Listening Assessment*

Text files representing annotated listening assessments of each speaker's performance, named according to speaker identification numbers, are contained in a single subdirectory. The text files were generated using WordPerfect 5.0. and 6.0., and are preserved in those formats, but have also been converted to ASCII. A sample file is included in this Handbook, as an appendix.

*8.2.3. Mistrack Files*

Text files representing mistrack intervals are contained in a single subdirectory, coded by speaker identification number. Post-processing operations will replace the coordinates of mistracked pellets with out-of-range values in the appropriate channels, and at appropriate times. However, even after post-processing, mistrack files may still provide useful information regarding certain causes of mistracks and their effects. These files will also provide sufficient information for verification of the treatment procedure for mistracks. A mistrack file, for one speaker, is included in the Handbook appendices.

*8.2.4. Speaker Anthropomorphics and Demographics*

A variety of ASCII text files, and white-space delimited ASCII tables, will be available that convey interesting and useful information about database speakers. Each file type will be restricted to a distinct subdirectory. Thus, subdirectories will exist for at least each of the following:

Dental characteristics (history, missing teeth, arch dimensions)
Cephalometric measures (cranial, mandibular, palatal [from outlines])

Pellet placements (tongue and mandible)
Demographics (age, gender, height, weight, education, foreign languages)
Residence (birth, formative, permanent)
Repaired records (corrected for synchronization flaws)

# REFERENCES

Abbs, J. H. (1989). *Speech research with an x-ray microbeam.* Unpublished NIDCD research grant proposal 5 P50 DC00162.

Abbs, J. H. & Nadler, R.D. (1987). *Users' manual for the University of Wisconsin X-ray Microbeam.* Madison, WI: University of Wisconsin Waisman Center.

Abbs, J. H. (1986). Invariance and variability in speech production: a distinction between linguistic intent and its neuromotor implementation. In J. S. Perkell and D. H. Klatt, (eds.), *Invariance and Variability in Speech Processes* (pp. 202-218). Hilldale, NJ: Lawrence Erlbaum Associates.

Angle, E. H. (1907). *Malocclusion of the Teeth.* Philadelphia: SS White Dental Manufacturing Co.

Behrents, R. G. (1985a). *Growth in the Aging Craniofacial Skeleton.* Ann Arbor, MI: Center for Human Growth and Development, The University of Michgan.

Behrents, R. G. (1985b). *An Atlas of Growth in the Aging Craniofacial Skeleton.* Ann Arbor, MI: Center for Human Growth and Development, The University of Michigan.

Broadbent, B. H., Sr., Broadbent, B. H., Jr., and Golden, W. H. (1975). *Bolton Standards of Dentofacial Developmental Growth.* St. Louis: C.V. Mosby.

Carterette, E. C. & Jones, M. H. (1974). *Informal Speech: Alphabetic and Phonemic Texts with Statistical Analyses and Tables.* Berkeley, CA: University of California Press.

Crystal, T. H. & House, A. S. (1982). Segmental durations in connected speech signals: preliminary results. *J. Acoust. Soc. Am., 72,* 705-716.

Crystal, T. H. & House, A. S. (1988a). The duration of American-English vowels: an overview. *J. Phonetics, 16,* 263-284.

Crystal, T. H. & House, A. S. (1988b). The duration of American-English stop consonants: an overview. *J. Phonetics, 16,* 285-294.

Crystal, T. H. & House, A. S. (1988c). Segmental durations in connected speech signals: current results. *J. Acoust. Soc. Am., 83,* 1553-1573.

Crystal, T. H. & House, A. S. (1988d). Segmental durations in connected speech signals: syllabic stress. *J. Acoust. Soc. Am., 83,* 1574-1585.

Darley, F. L., Aronson, A. E. & Brown, J. R. (1975). *Motor Speech Disorders.* Philadelphia: W. B. Saunders Co.

Edwards, J. & Harris, K. S. (1990). Rotation and transition of the jaw during speech. *J. Speech and Hearing Research, 33,* 550-562.

Elman, J. L. & Zipser, D. (1988). Learning the hidden structure of speech. *J. Acoust. Soc. Am., 83,* 1615-1626.

Farkas, L. G. (1981). *Anthropometry of the Head and Face in Medicine*. New York: Elsevier.

Fisher, W. M., Doddington, G. R. & Goudie-Marshalla, K. M. (1986). The DARPA speech recognition research database: specifications and status. In *Proceedings of the DARPA Speech Recognition Workshops* (pp. 93-99). Washington, DC: Science Applications International Corp.

Fujimura, O., Kiritani, S. & Ishida, H. (1973). Computer-controlled radiography for observation of movements of articulatory and other human organs. *Comput. Biol. Med., 3,* 371-384.

Fujimura, O. (ed.) (1988). Articulatory organization--from phonology to speech signals. *Phonetica, 45,* 77-197.

Garofolo, J. S. (1988). The structure and format of the DARPA TIMIT CD-ROM prototype. Manuscript distributed with the DARPA TIMIT Acoustic Phonetic Cintinuous Speech Database, produced by National Institute of Standards & Technology.

Gibbs, C. H. & Messerman, T. (1972). Jaw motion during speech. In R. T. Wertz (ed.), *Orofacial Function: Clinical Research in Dentistry and Speech Pathology* (pp. 104-112). Ann Arbor, MI: ASHA report # 7.

Goldstein, U. (1980). *An Articulatory Model for the Vocal Tracts of Growing Children*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Hultzen, L. S., Allen, J. H. D. & Miron, M. S. (1964). *Tables of Transitional Frequencies of English Phonemes*. Urbana, IL: University of Illinois Press.

Kent, R. D. (1972). Some considerations in the cinefluorographic analysis of tongue movement during speech. *Phonetica, 26, 16-32.*

Kiritani, S. (1986). X-ray microbeam method for measurement of articulatory dynamics: techniques and results. *Speech Communication, 5, 119-140.*

Klatt, D. H. (1986). The problem of variability in speech recognition and in models of speech perception. In J. S. Perkell and D. H. Klatt, (eds.), *Invariance and Variability in Speech Processes* (pp. 300-319). Hilldale, NJ: Lawrence Erlbaum Associates.

Klatt, D. H. (1987). Review of text to speech conversion for English. *J. Acoust. Soc. Am., 82,* 737-793.

Kucera, H. & Francis, W. N. (1967). *Computational Analysis of Present-Day American English*. Providence, RI: Brown University Press.

Kuehn, D. P. & Moll, K. L. (1976). A cineradiographic study of vc & cv articulatory velocities. *J. Phonetics, 4,* 303-320.

Lamel, L. F., Kassel, R. H. & Seneff, S. (1986). Speech database development: design and analysis of the acoustic-phonetic corpus. In *Proceedings of the DARPA Speech Recognition Workshops* (pp. 100-109). Washington, DC: Science Applications International Corp.

LeSavoy, M. A. (ed.) (1981). *Reconstruction of the Head and Neck* (pp. 2-9). Baltimore, MD: Williams and Wilkins.

Lindau, M. & Ladefoged, P. (1988). Pellet placement in x-ray microbeam study. *J. Acoust. Soc. Am., 84,* S1, 126.

Lubker, J. & Gay, T. (1982). Anticipatory labial coarticulation: experimental, biological , and linguistic variables. *J. Acoust. Soc. Am., 71,* 437-448.

MacWhinney, B., & Snow, C. E. (1985). The child language data exchange system, *J. Child Lang., 12,* 271-296.

Maddieson, I. (1984). *Patterns of Sounds*. Cambridge: Cambridge University Press.

Martin, R. E. (1991). *A Comparison of Lingual Movement in Swallowing and Speech Production*. Unpublished doctoral dissertation, University of Wisconsin-Madison.

McDonald, E. T. (1964). *A Deep Test of Articulation: Sentence Form*. Pittsburgh, PA: Stanwix House.

Ohala, J. J. (ed.) (1989). Theme issue: on the quantal nature of speech. *J. Phonetics, 17,* 1-157.

Öhman, S. E. G. & Stevens, K. N. (1963). Cineradiographic studies of speech: procedures and objectives. *J. Acoust. Soc. Am., 35,* 1889(A).

Papcun, G., Hochberg, J., Thomas, T. R., Laroche, F., Zacks, J. & Levy, S. (1992). Inferring articulation & recognizing gestures from acoustics with a neural network trained on x-ray microbeam data. *J. Acoust. Soc. Am., 92 ,* 688-700.

Perkell, J. S. (1986). On sources of invariance & variability in speech production. In J. S. Perkell and D. H. Klatt, (eds), *Invariance and Variability in Speech Processes* (pp. 260-263). Hilldale, NJ: Lawrence Erlbaum Associates.

Perkell, J. S. (1969). *Physiology of Speech Production*. MIT Press: Cambridge.

Peterson, G. E. & Barney, H. E. (1952). Control methods used in a study of vowels. *J. Acoust. Soc. Am., 24,* 693-703.

Pisoni, D. (1986). Contextual variability & the problem of acoustic-phonetic invariance in speech. In J. S. Perkell and D. H. Klatt, (eds.), *Invariance and Variability in Speech Processes* (pp. 154-160). Hilldale, NJ: Lawrence Erlbaum Associates.

Rodman, K., Moody, T., & Price, J. (1984). Speech recognizer performance of voice handicapped users. *Proc. AVIOS '84 Voice I/O System Applications Conf.*

Roget P.M. (1947). *Roget's Thesaurus of Words and Phrases*. (Preface to the First Edition [1852], p. ix). New York, NY: Grosset & Dunlap.

Rosenbek, J. C. & LaPointe, L. L. (1985). The dysarthrias: description, diagnosis, & treatment. In D. F. Johns, (ed.), *Clinical Management of Neurogenic Communicative Disorders* (pp. 97-152). Boston: Little, Brown and Co.

Short M. J. (1987). *Essential Anatomies: Oral & Head/Neck* (pp. 37-41; 154-160). Albany, NY: Delmar Publishers Inc.

Templin, M. C. & Darley, F. L. (1960). *The Templin Darley Tests of Articulation*. Iowa City, IA: Bureau of Educational Research and Service Extension Division, State University of Iowa.

Thompson, Murray (1984). *X-Ray Microbeam manual* (unpublished draft). Stoughton, WI: University of Wisconsin Physical Sciences Laboratory.

Umeda, N. (1977). Consonant duration in American English. *J. Acoust. Soc. Am., 61,* 846-856.

Umeda, N. (1975). Vowel duration in American English. *J. Acoust. Soc. Am., 58,* 434-445.

Vatikiotis-Bateson, E. & Ostry, D. (submitted). An analysis of the dimensionality of jaw motion in speech. *J. Phonetics*.

Weismer, G., Kent, R. D., Hodge, M. & Martin, R. (1988). The acoustic signature for intelligibility test words. *J. Acoust. Soc. Am., 84,* 1281-1291.

Westbury, J. R. (1991). The significance and measurement of head position during speech production experiments using the x-ray microbeam system. *J. Acoust. Soc. Am., 89,* 1782-1791.

Westbury, J. R. (1994). On coordinate systems and the representation of articulatory movements. *J. Acoust. Soc. Am., 95,* 2271-2273.

Yorkston, K. M. & Beukelman, D. R. (1981). *Assessment of Intelligibility of Dysarthric Speech*. Tigard, OR: C. C. Publications, Inc.

X-Ray Microbeam Speech Production Database

Task List

Task #1:    **words0l#001** 7500 msec.

CITATION WORDS

Read each word once.
Pause briefly between words.

problem
children
dormer
never
dormitory
school
has

Task **#2:**    **words02#002** 7500 msec.

CITATION WORDS

nothing
this
street
even
special
children
ship

Task **#3:**    **nseqsl#003** 10500 msec.

Phrases made fromNUMBERSEQUENCES:

Read each sequence once, as though it were
a sentence of seven words.  Pause briefly
between sequences.

9739286 8495571 5945341

Task **#4:**    **words03#004** 7500 msec.

CITATION WORDS

row
special
but
special
things
although
glowing

# Database Task List

Task **#5:** **words04#005** 7500 msec.

## CITATION WORDS

people
told
look
moment
programmer
moment
quite

Task **#6: words05#006** 7500 msec.

## CITATION WORDS

hall
this
right
dormer
told
already
blend

Task **#7:** **sent0l#007** 10500 msec.

### SENTENCES

Read each sentence once, at a comfortable conversational rate. Pause briefly between sentences.

1. She is about two or three.
2. When can we go home?
3. Hispanic costumes are quite colorful.

Task **#8: words06#008** 7500 msec.

## CITATION WORDS

form
ship
back
almost
things
school
programmer

Task **#9: words07#009** 7500 msec.

## CITATION WORDS

order
row
shoot
used
right
nothing
been

Task #10: **sent02#010** 10500 msec.

### SENTENCES

1. The other one is too big.
2.Don't do Charlie's dirty dishes.
3.She had your dark suit in greasy wash
water all year.

Task #11:     **gfathr#011a** 25000 msec.

READ THE PARAGRAPH
You wish to know all about my grandfather.
Well, he is nearly 93 years old, yet he still
thinks as swiftly as ever.  He dresses himself
in an old black frock coat, usually several
buttons missing.  A long beard clings to his
chin, giving those who observe him a
pronounced feeling of the utmost respect.
When he speaks, his voice quivers a bit.
Twice each day he plays skillfully upon a
small organ.

Task **#12:**     **gfathr#011b** 22000 msec.

READ THE PARAGRAPH
Twice each day he plays skillfully and with
zest upon a small organ.  Except in the
winter when the snow or ice prevents, he
slowly takes a short walk in the open air
each day.  We have often urged him to walk
more and smoke less, but he always
answers, 'Banana oil!' Grandfather likes to
be modern in his language.

Task **#13:**    **sVd's#012**     20000 msec.

CITATION svd's:

Read each item once, clearly,
with a brief pause between items.
(Read in column order.)

| side | sawed | *sud (dud) |
|------|-------|------------|
| sewed | sid | *soid (Lloyd) |
| seed | sad | *sowd (loud) |
| sod | surd | *sood (wood) |
| sued | said | *sayed (bayed) |

Task **#14:**    **vowels#013** 15000 msec.

CITATION VOWELS:

Read each item once, slowly and clearly,
with a brief pause between items.
(Read in colunm order.)

| er (dirt) | aw (bought) | ee (beet) |
|-----------|-------------|-----------|
| uh (but) | oh (boat) | eh (bet) |
| uu (boot) | ih (bit) | oo (foot) |
| ay (date) | ah (hot) | ae (bat) |

Task **#15:**    **vseqs#014**     10500 msec.

VOWEL SEQUENCES:

Read each item once, slowly and clearly,
with a brief pause between items.

[iu]  [ia]  [ua]  [au]  [ai]  [ui]
eeoo eeah ooah ahoo ahee ooee

Task **#16:**   **vcv's#015**   27000 msec.

CITATION VCV's
Read each item once, slowly and clearly,
with a brief pause between items.
(Read in colunm order.)

uhfA  uhzA  uhwA
uhkA  uhhA  uhgA
uhrA  uhchA  uhnA
uhmA  uhshA  uhpA
uhzhA  uhbA  uhdA
uhtA  uhyA  uhjA
uhvA  uhlA  uhsA

Task **#17:**   **sent03#016**   10500 msec.

SENTENCES

1.  Don't ask me to carry an oily rag
    like that.
2.  You can shoot at the ship or do nothing.
3.  You must blend certain things to make a
    special wax.

Task **#18:**   **words08#017** 7500 msec.

CITATION WORDS

dormitory
words
about
light
first
about
back

Task **#19:**   **sent04#018**   10500 msec.

SENTENCES

1.  The coat has a blend of both light
    and dark fibers.
2.  Across the street stands a country
    school.
3.  The dormitory is between the house
    and the school.

Task **#20:**   **sent05#019**   10500 msec.

SENTENCES

1.  I assume moisture will damage this
    ship's hull.
2.  The coat has a blend of both light
    and dark fibers.
3.  If I had that much cash I'd buy the
    house.

Task **#21:**   **words09#020** 7500 msec.

CITATION WORDS

country
understand
silk
sense
hall
both
country

# Database Task List

**Task #22:** **words10#021** 7500 msec.

### CITATION WORDS

school
dormer
children
seemed
house
had
but

**Task #23:** **words11#022** 7500 msec.

### CITATION WORDS

coat
blend
street
child
dormer
had
moment

**Task #24:** **sent06#023** 10500 msec.

### SENTENCES

1. Things in a row provide a sense of order.
2. Put these two back.
3. Second children are often special people.

**Task #25:** **words12#024** 7500 msec.

### CITATION WORDS

dorm
that
cash
through
children
going
making

**Task #26:** **slow1#025** 14000 msec.

### SLOW SPEAKING RATE

Repeat this pair of sentences twice, at HALF your normal speaking rate, without pausing between words. Pause briefly between sentences.

1. The other one is too big.
2. Combine all the ingredients in a large bowl.

.Task **#27:** **words13#026** 7500 msec.

### CITATION WORDS
seemed
yourself
across
right
sense
second
could

# Database Task List

Task **#28:  words14#027** 7500 msec.

CITATION WORDS

program
told
this
across
between
children
moment

Task **#29:     sent07#028**    10500 msec.

SENTENCES

1.  When all else fails, use force.
2.  You can shoot at the ship or do nothing.
3.  Put these two back.

Task **#30.   sent08#029**    10500 msec.

SENTENCES

1.  The point of the program will be told before long.
2.  Shaving cream is a popular item on Halloween.
3.  Put these two back.

Task **#31:     sent09#030**    10500 msec.

SENTENCES

1.  We are open every Monday evening.
2.  I'll make sense of the problem in a moment.
3.  They all know what I said.

Task **#32:  words15#031** 7500 msec.

CITATION WORDS

dorm
before
programmer
blend
sense
told
people

Task **#33:  words16#032** 7500 msec.

CITATION WORDS

beautiful
row
than
dorm
sense
second
things

**Database Task List**

Task **#34:**   **fast1#033**   10500 msec.

FAST SPEAKING RATE

Repeat this pair of sentences twice,
at 2 TIMES your normal speaking rate
without becoming unintelligible.  Pause
briefly between sentences.

1.   Combine all the ingredients in a large
     bowl.
2.   The other one is too big.

Task **#35:**       **words17#034** 7500 mscc.

CITATION WORDS

much
order
smooth
people
have
would
special

Task **#36:**   **sent10#035**     10500 msec.

SENTENCES

1.   They remained lifelong friends and
     companions.
2.   They all know what I said.
3.   The other one is too big.

Task **#37:**   **words18#036**   7500 msec.

CITATION WORDS

sense
long
house
across
programmer
problem
question

Task **#38:**   **slow2#037**     14000 msec.

SLOW SPEAKING RATE

Repeat this pair of sentences twice at
HALF your normal speaking rate,
without pausing between words.
Pause briefly between sentences.

1.   The other one is too big.
2.   Combine all the ingredients in a large
     bowl.

Task **#39:**   **sent11#038**     10500 msec.

SENTENCES

1.   When all else fails, use force.
2.   Combine all the ingredients in a
     large bowl.
3.   Things in a row provide a sense of
     order.

# Database Task List

Task **#40:**     **sent12#039**    10500 msec.

SENTENCES

1. You must blend certain things to make a special wax.
2. The dormitory is between the house and the school.
3. Don't ask me to carry an oily rag like that.

Task **#41:   words19#040** 7500 msec.

CITATION WORDS

but
point
about
ship
house
early
things

Task **#42:    fast2#041**    10500 msec.

FAST SPEAKING RATE

Repeat this pair of sentences twice, at 2 TIMES your normal speaking rate, without becoming unintelligible.  Pause briefly between sentences.

1. Combine all the ingredients in a large bowl.
2. The other one is too big.

Task **#43:    sent13#042**    10500 msec.

SENTENCES

1. You can shoot at the ship or do nothing.
2. The gorgeous butterfly ate a lot of nectar.
3. You must blend certain things to make a special wax.

Task **#44:       words20#043** 7500 msec.

CITATION WORDS

seemed
problem
moment
become
were
seemed
work

Task **#45:    sent14#044**    10500 msec.

SENTENCES

1. The other one is too big.
2. She always jokes about too much garlic in his food.
3. If I had that much cash I'd buy the house.

# Database Task List

Task **#46:** **sent15#045**  10500 msec.    .

### SENTENCES

1. The point of the program will be told before long.
2. Across the street stands a country school.
3. She had your dark suit in greasy wash water all year.

Task **#47:**  **words21#046** 7500 msec.

### CITATION WORDS

dormitory
among
second
street
across
find
row

Task **#48:**  **sentl6#047**   10500 msec.

### SENTENCES
1. The point of the program will be told before long.
2. Second children are often special people.
3. I'll make sense of the problem in a moment.

Task **#49:**  **words22#048** 7500 msec.

### CITATION WORDS

row
order
problem
country
told
dorm
coat

Task **#50:**  **sent17#049**   10500 msec.

### SENTENCES

1.  The other one is too big.
2.  They all know what I said.
3.  Things in a row provide a sense of order.

Task **#51: nseqs2#050**  10500 msec.

### NUMBER SEQUENCES

7789388  8761335  2918524

# Database Task List

**Task #52:** **words23#051** 7500 msec.

### CITATION WORDS

program
dorm
dormer
order
didn't
house
program

**Task #53:** **sent18#052** 10500 msec.

### SENTENCES

1. She had your dark suit in greasy wash water all year.
2. You can shoot at the ship or do nothing.
3. Combine all the ingredients in a large bowl.

**Task #54:** **words24#053** 7500 msec,

### CITATION WORDS

long
light
programmer
information
above
sigh
sip

**Task #55:** **sent19#054** 10500 msec.

### SENTENCES

1. Put these two back.
2. The point of the program will be told before long.
3. The coat has a blend of both light and dark fibers.

**Task #56:** **sent20#055** 10500 msec.

### SENTENCES

1. The sermon emphasized the need for affirmative action.
2. Second children are often special people.
3. That noise problem grows more annoying each day.

**Task #57:** **slow3#056** 14000 msec.

### SLOW SPEAKING RATE

Repeat this pair of sentences twice, at HALF your normal speaking rate, without pausing between words. Pause briefly between sentences.

1. Combine all the ingredients in a large bowl.
2. The other one is too big.

**Database Task List**

Task **#58:**    **words25#057**    7500 msec.

CITATION WORDS

around
both
country
had
ship
yet
both

Task **#59:**    **sent21#058**    14500 msec.

SENTENCES

1. The coat has a blend of both light
   and dark fibers.
2. They all know what I said.
3. I'll make sense of the problem in a
   moment.
4. Combine all the ingredients in a
   large bowl.

Task **#60:**    **sent22#059**    10500 msec.

SENTENCES

1. Things in a row provide a sense of order.
2. Put these two back.
3. She had your dark suit in greasy wash
   water all year.

Task **#61:**    **words26#060** 7500 msec.

CITATION WORDS

point
man
enjoy
long
much
shoot
had

Task **#62:**    **words27#061** 7500 msec

CITATION WORDS

against
people
first
long
from
people
weigh

Task **#63:**    **nseqs3#062**    10500 msec.

NUMBER SEQUENCES

6582269 7217424 2315483

# Database Task List

Task **#64:** **sent23#063** 10500 msec.

SENTENCES

1. I'll make sense of the problem in a moment.
2. The point of the program will be told before long.
3. Second children are often special people.

Task **#65:** **words28#064** 7500 msec.

CITATION WORDS

coat
began
cash
blend
this
pushed
flip

Task **#66:** **words29#065** 7500 msec.

CITATION WORDS

shoot
country
both
shoot
cash
program
second

Task **#67:** **sent24#066** 10500 msec.

SENTENCES

1. When all else fails, use force.
2. They all know what I said.
3. The oasis was a mirage.

Task **#68:** **sent25#067** 10500 msec.

SENTENCES

1. When all else fails, use force.
2. Don't ask me to carry an oily rag like that.
3. Across the street stands a country school.

Task **#69:** **sent26#068** 10500 msec.

SENTENCES

1. Things in a row provide a sense of order.
2. I'll make sense of the problem in a moment.
3. The coat has a blend of both light and dark fibers.

**Database Task List**

Task **#70:** **words3O#O69** 7500 msec.

CITATION WORDS

so
much
that
light
there
house
special

Task **#73:** **words31#072** 7500 msec.

CITATION WORDS

cash
nothing
point
what
school
that
because

Task **#71:** **sent27#070** 10500 msec.

SENTENCES

1. Do they go up and down?
2. You must blend certain things to make a special wax.
3. Across the street stands a country school.

Task **#74:** **sent28#073** 10500 msec.

SENTENCES

1. The dormitory is between the house and the school.
2. Combine all the ingredients in a large bowl.
3. Don't ask me to carry an oily rag like that.

Task **#72: nseqs4#071** 7000 msec.

NUMBER SEQUENCES

5681998      6744166

Task **#75:** **emphl#074** 17500 msec.

Repeat the question-answer pair 5 times.

EMPHASIZE the word that is corrected
and capitalized in the response.

Put those two back?
No, put THESE two back.

# Database Task List

Ta sk **#76:  words32#075**  7500 msec.

## CITATION WORDS

major
first
about
back
this
nothing
wax

Task **#77:   sent29#076**   10500 msec.

## SENTENCES

1. It's just a little thing.
2. Grandmother outgrew her upbringing in petticoats.
3. If I had that much cash I'd buy the house.

Task **#78:     hunter#077a** 25000 msec.

## READ THE PARAGRAPH

In late fall and early spring the short rays of the sun call a true son of the out-of-doors back to the places of his childhood.  Tom Brooks was such a man.  Each year at these times his desk seemed like a stone whose weight made him wish for the life he knew as a boy.  In the five years since leaving college he had not revisited his old haunts before.  But this March Tom found himself by a small stream with a gun.

Task **#79:     hunter#077b** 25000 msec.

## READ THE PARAGRAPH

This March, Tom found himself by a small stream with a gun at rest in the crook of his arm. The desk that had tied him down was gone and his one thought was for quail. He had been on the trail since dawn, but not one bird had crossed his path. It seemed as though five years without hunting had made him lose touch with all the small signs that he once knew - signs that would tell for sure if  an animal was near or not.

Task **#80:     hunter#077c** 22000 msec.

## READ THE PARAGRAPH

Once he thought he saw a bird, but it was just a large leaf that had failed to drop to the ground during the winter. Tom stopped near a small stream to rest. Soon after he had laid down his gun, he heard the sound of wings from across the stream, and five large birds came out of the brush.  They flew to the edge of the stream unaware of the hunter.

Task **#81:** **hunter#077d** 21000 msec.

READ THE PARAGRAPH

The birds flew to the edge of the stream
unaware of the hunter.  Tom placed his hand
on his gun quietly.  Slowly he raised it to his
shoulder and took aim.  The seconds ticked
off like hours, but still the birds drank.
Quick shots rang out.  The years of waiting
seemed to disappear with the successful
culmination of the hunt.

Task **#82:** **clear#078** 17500 msec

CLEAR SPEECH

 Repeat this sentence five times,
  VERY DISTINCTLY and CLEARLY,    as
if you are trying to make someone
understand you in a noisy environment.
   Do not pause between words.

Combine all the ingredients in a large
 bowl.

Task **#83:** **words33#079** 7500 msec.

CITATION WORDS

seemed
himself
point
hail
wax
dormitory
hail

Task **#84:** **sent30#080** 10500 msec.

SENTENCES

1.      You must blend certain things to
        make a special wax.
2.      1 think that's real.
3.      Combine all the ingredients in a
        large bowl.

Task **#85:** **sent31#081** 10500 msec.

SENTENCES

1.      Porcupines resemble sea urchins.
2.      Across the street stands a country
        school.
3.      Second children are often special
        people.

Task **#86:** **sent32#082** 10500 msec.

SENTENCES

1.      Put this one right here.
2.      Cheap stockings run the first time
        they're worn.
3.      The dormitory is between the house
        and the school.

**Database Task List**

Task **#87:**    **words34#083** 7500 msec.

CITATION WORDS

first
second
head
long
dormitory
that
right

Task **#88:**    **count#084**    20000 msec.

  COUNT from 1 to 20, clearly and at a
moderate rate, with brief pauses between
numbers. (Do NOT try to do it on one
breath.)

Task **#89:**    **words35#085** 7500 msec.

CITATION WORDS

himself
cash
point
conversation
that
shoot
had

Task **#90:**    **words36#086** 7500 msec.

CITATION WORDS

but
zero
much
program
himself
both
across

Task **#91:**    **words37#087** 7500 msec.

CITATION WORDS

hail
light
wax
light
ship
blend
school

Task **#92:**    **nseqs5#088** 10500 msec.

NUMBER SEQUENCES

4375125      3647962      1146327

Task **#93:**    **fast3#089**    10500 ms

### FAST SPEAKING RATE

  Repeat this pair of sentences twice, at 2 TIMES your normal speaking rate, without becoming unintelligible.    Pause briefly between sentences.

1.   The other one is too big.
2.   Combine all the ingredients in a large bowl.


Task **#94:**    **emph2#090** 17500 msec.

Repeat the question-answer pair 5 times.

EMPHASIZE the word that is corrected and capitalized in the response.

Put these two down?
No, put these two BACK


Task **#95:   words38#091**   7500 msec.

### CITATION WORDS

himself
blink
but
about
measure
coat
wax


Task **#96:**    **sent33#092**    10500 msec.

### SENTENCES

1.    Does Creole cooking use curry?
2.    A roll of wire lay near the wall.
3.    Don't ask me to carry an oily rag like that.


Task **#97:**    **sent34#093**    10500 msec.

### SENTENCES

1.    The other one is too big.
2.    If I had that much cash I'd buy the house.
3.    You can shoot at the ship or do nothing.


Task **#98:**    **sent35#094**    10500 msec.
.
### SENTENCES

1. If I had that much cash I'd buy the house.
2. She had your dark suit in greasy wash water all year.
3. Put these two back.

**Database Task List**

Task **#99:**      **words39#095** 7500 msec.

CITATION WORDS

wax
order
coat
right
problem
first
much

Task #**100**:    **words40#096** 7500 msec.

CITATION WORDS

himself
back
street
nothing
back
things
street

Task #101:    **sent36#097**     10500 msec.

SENTENCES

1. Elderly people are often excluded.
2. When all else fails, use force.
3. The dormitory is between the house and the school.

Task **#102: diadop#098** 3000 msec.

Diadochokinesis:
Repeat as fast as possible

puh,puh,puh,puh ...

Task **#103: diadot#099**    3000 msec.

Diadochokinesis:
Repeat as fast as possible

tuh, tuh, tuh, tuh ...

Task **#104: diadok#100** 3000 msec.

Diadochokinesis:
Repeat as fast as possible

kuh,kuh,kuh,kuh ...

Task **#105: sa#101**     10500 msec.

Large-amplitude, repeated [sa]

sa, sa, sa ...

Task **#106: wag#102**    10500 msec.
Large-amplitude, repeated jaw-wagging

silently, (wag), (wag), (wag) ...

# Database Task List

Task **#107:**   **swal2l#103**  3000 msec.

Water Swallow: 2cc

Task **#108:**   **swal22#104**  3000 msec.

Water Swallow: 2cc

Task **#109:**   **swal23#105**  3000 msec.

Water Swallow: 2cc

Task #110:   **swal24#106**  3000 msec.

Water Swallow: 2cc

Task #**111**:   **swal25#107** 3000 msec.

Water Swallow: 10cc

Task #**112**:   **swal25#107** 3000 msec.

Water Swallow:  10cc

Task **#113:**   **swal102#109** 3000 msec.

Water Swallow: 10cc

Task **#114:**   **swal103#110** 3000 msec.

Water Swallow: 10cc

Task **#115:**   **swal104#111** 3000 msec.

Water Swallow: 10cc

Task **#116:**   **swal105#112**

Water Swallow: 10cc

Task **#117: max#113**    3500 msec.

MAXIMUM TONGUE PROTRUSION

Stick your tongue out of your mouth
as far as possible.

Task **#118:   max#114**      3500 msec.

MAXIMUM LIP PROTRUSION

Purse your lips forward
as far as possible.

**APPENDIX B:** Accumulated records, by task, across all Database speakers.

Note: While the total number of different speakers is 57, there are often more than 57 exemplars of a given task, since a task may have been repeated for some speakers, for any number of reasons (see Handbook text). Conversely, various data collection or processing difficulties may have resulted in the loss of some target records, and a grand total less than 57.

| | | | |
|---|---|---|---|
| clear speech (52) | sent15 (54) | words01 (60) | words40 (51) |
| count (58) | sent16 (54) | words02 (59) | |
| diadok (58) | sent17 (54) | words03 (60) | |
| diadop (65) | sent18 (54) | words04 (58) | |
| diadot (58) | sent19 (53) | words05 (59) | |
| emphasis1 (50) | sent20 (53) | words06 (59) | |
| emphasis2 (53) | sent21 (52) | words07 (58) | |
| fast1 (53) | sent22 (52) | words08 (55) | |
| fast2 (52) | sent23 (50) | words09 (54) | |
| fast3 (49) | sent24 (55) | words10 (56) | |
| grndfather[a] (57) | sent25 (52) | words11 (56) | |
| grndfather[b] (53) | sent26 (51) | words12 (56) | |
| hunter[a] (50) | sent27 (51) | words13 (56) | |
| hunter[b] (46) | sent28 (50) | words14 (54) | |
| hunter[c] (46) | sent29 (51) | words15 (55) | |
| hunter[d] (45) | sent30 (51) | words16 (54) | |
| max#113 (76) | sent31 (51) | words17 (54) | |
| max#114 (51) | sent32 (51) | words18 (53) | |
| nseqs1 (67) | sent33 (51) | words19 (54) | |
| nseqs2 (60) | sent34 (51) | words20 (54) | |
| nseqs3 (53) | sent35 (53) | words21 (54) | |
| nseqs4 (63) | sent36 (51) | words22 (54) | |
| nseqs5 (58) | slow1 (54) | words23 (54) | |
| sVd's (62) | slow2 (51) | words24 (55) | |
| sa (65) | slow3 (48) | words25 (52) | |
| sent01 (57) | swal101 (52) | words26 (51) | |
| sent02 (57) | swal102 (53) | words27 (50) | |
| sent03 (54) | swal103 (49) | words28 (51) | |
| sent04 (56) | swal104 (51) | words29 (51) | |
| sent05 (56) | swal105 (52) | words30 (51) | |
| sent06 (56) | swal21 (52) | words31 (51) | |
| sent07 (55) | swal22 (50) | words32 (51) | |
| sent08 (54) | swal23 (50) | words33 (52) | |
| sent09 (56) | swal24 (51) | words34 (51) | |
| sent10 (55) | swal25 (53) | words35 (51) | |
| sent11 (54) | vcv's (60) | words36 (50) | |
| sent12 (56) | vowels (69) | words37 (51) | |
| sent13 (54) | vowseqs (67) | words38 (51) | |
| sent14 (55) | wag (58) | words39 (51) | |

**Appendix C**:  Phonetic Transcription of Speaking Task

WORDS O1
P R AA B L AX M
CH IH L D R AX N
D AO R M AXR
N EH V AXR
D AO R M AX T AO R IY
S K UX L
H AE Z

WORDS O2
N AH TH IH NG
DH IH S
S T R IY Q T
IY V AX N
S P EH SH AX L
CH IH L D R AX N
SH IH P

NSEQS 01
N AY N * S EH V AX N * TH R IY * N AY N * T UX * EY Q T * S IH K S/
EY Q T * F AO R * N AY N * F AY V * F AY V * S EH V AX N * S EH V AX N/
F AY V * N AY N * F AO R * F AY V * TH R IY * F AO R * W AH N

WORDS 03
R OW
S P EH SH AX L
B AH Q T
S P EH SH AX L
TH IH NG Z
AO L DH OW
G L OW IH NG

WORDS 04
P IY P AX L
T OW L D
L UH K
M OW M EH N T
K W AY Q T

WORDS 05
H EY L
DH IH S
R AY Q T
D AO R M AXR

T OW L D
AO L R EH Q IY
B L EH N D

1.  SH IY * IH Z * AX B AW Q T * T UX * AXR * TH R IY
2.  W EH N * K EH N * W IY * G OW * H OW M
3.  H IH S P AE N IH Q K * K AA S T UX M Z * AA R * K W AY Q T * K AH L AXR F AX
L


WORDS 06
F AO R M
SH IH P
B AE Q K
AO L M OW S T
TH IH NG Z
S K UX L
P R OW G R AE M AX


WORDS 07
AO R D AXR
R OW
SH UX Q T
Y UX Z D
R AY Q T
N AH TH IH NG
B EH N


SENT 02
1.  DH IY * AH DH AX * W AH N * IH Z * T UX * B IH G
2.  D OW N Q T * D UX * CH AA R L IY Z * D ER DX IY * D IH SH AX Z
3.  SH IY * H AE D * Y AXR * D AA R Q K * S UX Q T * AX N * G R IY S IY * W AA
SH * W AA DX AXR * AO L * Y IH R


GFATHR 11A
Y UX * W IH SH * T AX * N OW * AO L * AX B OW Q T * M AY * G R AE N F AA DH
AXR */ W EH L * HH IY * AX Z * N IH R L IY * N AY N DX IY * TH R IY * Y IH R Z *
OW L D * Y EH DX IY * S T IH L * TH IH NG K S * AE Z * S W IH F D L IY * AE Z * EH
V AXR */ HH IY * D R EH S AX Z * HH IH M S EH L F * IH N * AX N * OW L D * B L AE
Q K * F R AA Q K * K OW Q T * Y UX ZH UX AX L IY * S EH V R AX L * B AH Q N Z *
M IH S IH NG */ AH * L AO NG * B IH R D * K L IH NG Z * T UX * HH IH Z * CH IH N *
G IH V IH NG * DH OW Z * HH UX * AX B Z ER V * HH IH M * AH * P R AX N AW N S
T * F IY L IH NG * AH V * DH AX * AH DX M OW S Q T * R AX S P EH K T */ W EH N *
HH IY * S P IY K S * HH IH Z * V OY S * K W IH V AXR Z * AH * B IH Q T */ T W AY S *
IY CH * D EY * HH IY * P L EY Z * S K IH L F AX L IY * AH P AA N * AH * S M AO L *
AO R G AX N

GFATHR 11B

T W AY S * IY CH * D EY * HH IY * P L EY Z * S K IH L F AX L IY * AE N * W IH
TH * Z EH S T * AH P AA N * AH * S M AO L * AO R G AX N */ EH K S EH P T * IH N *
DH AH * W IH N T AXR * W EH N * DH AH * S N OW * AO R * AY S * P R AX V EH N T
S * HH IY * S L OW L IY * T EY K S * AH * SH AO R Q T * W AO Q K * IH N * DH AH *
OW P AX N * EH R * IY CH * D EY */ W IY * HH AE V * AO F AX N * ER JH D * HH IH
M * T UX * W AO Q K * M OA R * AE N * S M OW K * L EH S * B AH Q T * HH IY * AO
L W EY Z * AE N S AXR Z * B AH N AE N AX * OY L */ G R AE N * F AA DH AXR * L
AY Q K S * T AX * B IY * M AA D AXR N * IH N AX Z * L AE NG G W IH JH

SVD'S

| | | |
|---|---|---|
| S AY D | S AO D | S AH D |
| S OW D | S IH D | S OY D |
| S IY D | S AE D | S AW D |
| S AA D | S ER D | S UH D |
| S UX D | S EH D | S EY D |

VOWELS

| | | |
|---|---|---|
| ER | AO | IY |
| AH | OW | EH |
| UX | IH | UH |
| EY | AA | AE |

VSEQS

IY UX * IY AA * UX AA * AA UX * AA IY * UX IY

VCV'S

| | | |
|---|---|---|
| AH F AH | AH Z AH | AH W AH |
| AH K AH | AH HH AH | AH G AH |
| AH R AH | AH CH AH | AH N AH |
| AH M AH | AH SH AH | AH P AH |
| AH ZH AH | AH B AH | AH D AH |
| AH T AH | AH Y AH | AH JH AH |
| AH V AH | AH L AH | AH S AH |

SENT 3

1. D OW N Q T * AE S K * M IY * DX AX * K AE R IY * AX N * OY L IY * R AE G * L
AY Q K * DH AE Q T
2. Y UX * K AX N * SH UX DX * AE Q T * DH AX * SH IH P * AXR * D UX * N AH TH
IH NG
3. Y UX * M AH S T * B L EH N D * S ER Q AX N * TH IH NG Z * T AX * M EY K * AX *
S P EH SH AX L * W AE K S

WORDS 8

D AO R M AX T AO R IY
W ER Q Z

AX B AW Q T
L AY Q T
F ER S T
AX B AW Q T
B AE Q K

SENT 4
1.  DH AX * K OW Q T * HH AE Z * AX * B L EH N D * AX V * B OW TH * L AY Q T *
AX N * D AA R K * F AY B AXR Z
2.  AH K R AO S * DH AX * S T R IY Q T * S T AE N Z * AX * K AH N T R IY * S K
UX L
3.  DH AX * D AO R M AX T AO R IY * IH Z * B AX T W IY N * DH AX * HH AW S
* AX N * DH AX * S K UX L

SENT 5
1.  AY * AH S UX M * M OY S CH AXR * W IH L * D AE M EH JH * DH IH S * SH IH P S
* HH AH L
2.  DH AX * K OW Q T * HH AE Z * AX * B L EH N D * AX V * B OW TH * L AY Q
T * AX N * D AA R K * F AY B AXR Z
3.  IH F * AY * HH AE D * DH AE Q T * M AH CH * K AE SH * AY D * B AY * DH AX *
HH AW S

WORDS 9
K AH N T R IY
AH N D AXR S T AE N D
S IH L K
S EH N S
HH EY L
B OW TH
K AH N T R IY

WORDS 10
S K UX L
D AO R M AXR
CH IH L D R AX N
S IY M D
H AW S
HH AE D

WORDS 11
K OW Q T
B L EH N D
S T R IY Q T
CH AY L Q D
D AO R M AXR
HH AE D

M OW M EH N Q T

1.  TH IH NG Z * IH N * AX * R OW * P R AX V AY D * AX * S EH N S * AX V * AO R D
AXR
2.  P UH Q T * DH IY Z * T UX * B AE Q K
3.  S EH K AX N * CH IH L D R AX N * AXR * AO F AX N * S P EH SH AX L * P IY
P AX L

D AO R M
DH AE T
K AE SH
TH R UX
CH IH L D R AX N
G OW IH NG
M EY K IH NG

1.  DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G
2.  K AH M B AY N * AO L * DH IY * IH N G R IY D IY AX N Q T S * IH N * AX * L
AA R JH * B OW L

S IY M D
Y AO R S EH L F
AX K R AO S
R AY Q T
S EH N S
S EH K AX N Q D
K UX Q D

P R OW G R AE M
T OW L Q D
DH IH S
AX D R AO S
B AX T W IY N
CH IH L D R AX N
M OW M EH N Q T

1.  W EH N * AO L * EH L S * F EY L Z * Y UX S * F AO R S
2.  Y UX * K AX N * SH UX Q T * AX Q T * DH AX * SH IH P * AXR * D UX * N AH TH I
NG
3.  P UH Q T * DH IY Z * T UX * B AE Q K

SENT 8
1.  DH AX * P OY N T * AX V * DH AX * P R OW G R AE M * W IH L * B IY * T OW
L Q D * B AX F AO R * L AO NG
2.  SH EY V IH NG * K R IY M * IH Z * AX * P AA P Y AX L AXR * AY DX AX M *
AA N * HH AA L AX W IY N
3.  P UH Q T * DH IY Z * T UX * B AE Q K

SENT 9
1.  W IY * AA R * OW P AX N * EH V R IY * M AH N D EY * IY V N IH NG
2.  AY L * M EY Q K * S EH N S * AX V * DH AX * P R AA B L AX M * IH N * AX *
M OW M AX N Q T
3.  DH EY * AO L * N OW * W AH DX * AY * S EH D

WORDS 15
D AO R M
B IY F AO R
P R OW G R AE M AXR
B L EH N D
S EH N S
T OW L Q D
P IY P AX L

WORDS 16
B Y UX T AX F AH L
R OW
DH AE N
D AO R M
S EH N S
S EH K AX N D
TH IH NG Z

FAST 1
1.  K AX M B AY N * AO L * DH AX * IH N G R IY D IY AX N T S * IH N * AX * L
AA R JH * B OW L
2.  DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G

WORDS 17
M AH CH
AO R D AXR
S M UX DH
P IY P AX L
HH AE V
W UH D
S P EH SH AX L

SENT 10

1. DH EY * R AX M EY N D * L AY F * L AO NG * F R EH N Z * AX N * K AX M P
AE N Y AX N Z
2. DH EY * AO L * N OW * W AH DX * AY * S EH D
3. DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G

WORDS 18
S EH N S
L AO NG
HH AW S
AX K R AO S
P R OW G R AE M AXR
P R AA B L EH M
K W EH S CH Y AX N

SLOW 2
1. DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G
2. K AX M B AY N * AO L * DH AX * IH N G R IY D IY AX N T S * IH N * AX * L
AA R JH * B OW L

SENT 11
1. W EH N * AO L * EH L S * F EY L Z * Y UX Z * F AO R S
2. K AX M B AY N * AO L * DH AX * IH N G R IY D IY AX N T S * IH N * AX * L
AA R JH * B OW L
3. TH IH NG S * IH N * AX * R OW * P R AX V AY D * AX * S EH N S * AX V * AO R D
AXR

SENT 12
1. Y UX * M AH S T * B L EH N D * S ER DX AX N * TH IH NG Z * T UX * M EY K
* AX * S P EH SH AX L * W AE K S
2. DH AX * D AO R M AX T AO R IY * IH Z * B AX T W IY N * DH AX * H AW S * AX N
* DH AX * S K UX L
3. D OW N Q T * AE S K * M IY * DX AX * K AE R IY * AX N * OY L IY * R AE G *
L AY Q K * DH AE Q T

WORDS 19
B AH Q T
P OY N T
AX B AW Q T
SH IH P
HH AW S
ER L IY
TH IH NG Z

FAST 2
1. K AX M B AY N * AO L * DH AX * IH N G R IY D IY AX N T S * IH N * AX * L
AA R JH * B OW L

2.  DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G


SENT 13
1. Y UX * K AX N * SH UX DX * AE Q T * DH AX * SH IH P * ER * D UX * N AH TH IH
NG
2.  DH AX * G AO R JH AX S * B AH DX AXR F L AY * EY DX * AX * L AA DX * AX V *
N EH K T AXR
3.  Y UX * M AH S T * B L EH N D * S ER DX AX N * TH IH NG Z * T UX * M EY K
* AX * S P EH SH AX L * W AE K S


WORDS 20
S IY M D
P R AA B L EH M
M OW M EH N Q T
B IY K AH M
W ER
S IY M D
W ER K


SENT 14
1.  DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G
2.  SH IY * AO L W AX Z * JH OW K S * AX B AW Q * T UX * M AH CH * G AA R L
AX K * IH N * HH IH Z * F UX D
3.  IH F * AY * HH AE D * DH AE Q T * M AH CH * K AE SH * AY D * B AY * DH AX *
HH AW S


SENT 15
1.  DH AX * P OY N T * AX V * DH AX * P R OW G R AE M * W IH L * B IY * T OW
L D * B AX F AO R * L AO NG
2.  AX K R AO S * DH AX * S T R IY Q T * S T AE N Z * AX * K AH N T R IY * S K
UX L
3.  SH IY * HH AE D * Y ER * D AA R K * S UX Q T * AX N * G R IY S IY * W AA SH *
W AA DX AXR * AO L * Y IY R


WORDS 21
D AO R M AX T AO R IY
AX M AH NG
S EH K AX N D
S T R IY Q T
AX K R AO S
F AY N Q D
R OW


SENT 16

112

1. DH AX * P OY N Q T * AX V * DH AX * P R OW G R AE M * W IH L * B IY * T OW L D * B AX F AO R * L AO NG
2. S EH K AX N * CH IH L D R AX N * AA R * AO F AX N * S P EH SH AX L * P IY P AX L
3. AY L * M EY Q K * S EH N S * AX V * DH AX * P R AA B L AX M * IH N * AX * M OW M AX N Q T

WORDS 22
R OW
AO R D AXR
P R AA B L EH M
K AH N T R IY
T OW L Q D
D AO R M
K OW Q T

SENT 17
1. DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G
2. DH EY * AO L * K OW * W AH DX * AY * S EH D
3. TH IH NG Z * IH N * AX * R OW * P R OW V AY D * AX * S EH N S * AX V * AO R D AXR

NSEQS 2
1. S EH V AX N * S EH V AX N * EY Q T * N AY N * TH R IY * EY Q T * EY Q T
2. EY Q T * S EH V AX N * S IH K S * W AH N * TH R IY * TH R IY * F AY V
3. T UX * N AY N * W AH N * EY Q T * F AY V * T UX * F AO R

WORDS 23
P R OW G R AE M
D AO R M
D AO R M AXR
AO R D AXR
D IH DX N T
HH AW S
P R OW G R AE M

SENT 18
1. SH IY * HH AE D * Y ER * D AA R K * S UX Q T * AX N * G R IY S IY * W AA SH * W AA DX AXR * AO L * Y IY R
2. Y UX * K AX N * SH UX DX * AX Q T * DH AX * SH IH P * AO R * D UX * N AH TH IH NG
3. K AX M B AY N * AO L * DH IY * IH N G R IY D IY AX N T S * IH N * AX * L AA R JH * B OW L

WORDS 24
L AO NG

L AY Q T
P R OW G R AE M AXR
IH N F AO R M EY SH AX N
AX B AH V
S AY
S IH P


SENT 19
1.  P UH Q T * DH IY Z * T UX * B AE Q K
2.  DH AX * P OY N T * AX V * DH AX * P R OW G R AE M * W IH L * B IY * T OW
L D * B AX F AO R * L AO NG
3.  DH AX * K OW Q T * HH AE Z * AX * B L EH N D * AX V * B OW TH * L AY Q
T * AX N * D AA R K * F AY B AXR Z


SENT 20
1.  DH AX * S ER M AX N * EH M F AX S AY Z D * DH AX * N IY D * F AO R * AX
F ER M AX DX AX V * AE K SH AX N
2.  S EH K AX N * CH IH L D R AX N * AA R * AO F AX N * S P EH SH AX L * P IY
P AX L
3.  DH AE Q T * N OY Z * P R AA B L AX M * G R OW Z * M AO R * AX N OY IH NG *
IY CH * D EY


SLOW 3
1.  K AX M B AY N * AO L * DH AX * IH N G R IY D IY AX N T S * IH N * AX * L
AA R JH * B OW L
2.  DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G


WORDS 25
AX R AW N D
B OW TH
K AH N T R IY
HH AE D
SH IH P
Y EH Q T
B OW TH


SENT 21
1.  DH AX * K OW Q T * HH AE Z * AX * B L EH N D * AX V * B OW TH * L AY Q
T * AX N * D AA R K * F AY B AXR Z
2.  DH EY * AO L * N OW * W AH DX * AY * S EH D
3.  AY L * M EY K * S EH N S * AX V * DH AX * P R AA B L AX M * IH N * AX * M
OW M AX N Q T
4.  K AX M B AY N * AO L * DH AX * IH N G R IY D IY AX N T S * IH N * AX * L
AA R JH * B OW L


SENT 22

1.  TH IH NG S * IH N * AX * R OW * P R OW V AY DX * AX * S EH N S * AX V *
AO R D AXR
2.  P UH Q T * TH IY Z T UX * B AE Q K
3.  SH IY * HH AE D * Y ER * D AA R K * S UX Q T * AX N * G R IY S IY * W AA SH *
W AA DX AXR * AO L * Y IY R

WORDS 26
P OY N Q T
M AE N
AX N JH OY
L AO NG
M AH CH
SH UX Q T
HH AE D

WORDS 27
AH G AX N S T
P IY P AX L
F ER S T
L AO NG
F R AH M
P IY P AX L
W EY

NSEQS 3
1.  S IH K S * F AY V * EY Q * T UX * T UX * S IH K S * N AY N
2.  S EH V AX N * T UX * W AH N * S EH V AX N * F AO R * T UX * F AO R
3.  T UX * TH R IY * W AH N * F AY V * F AO R * EY Q T * TH R IY

SENT 23
1.  AY L * M EY K * S EH N S * AX V * DH AX * P R AA P L AX M * IH N * AX * M
OW M AX N Q T
2.  DH AX * P OY N DX * AX V * DH AX * P R OW G R AE M * W IH L * B IY * T OW L
D * B AX F AO R * L AO NG
3.  S EH K AX N * CH IH L D R AX N * AA R * AO F AX N * S P EH SH AX L * P IY
P AX L

WORDS 28
K OW Q T
B IY G AE N
K AE SH
B L EH N D
P UH SH Q D
F L IH P

WORDS 29

SH UX Q T
K AH N T R IY
B OW TH
K AE SH
P R OW G R AE M
S EH K AX N D

1.  W EH N * AO L * EH L S * F EY AX L Z * Y UX Z * F AO R S
2.  DH EY * AO L * N OW * W AH DX * AY * S EH D
3.  DH AX * OW EY S IH S * W AH Z * AX * M ER AA ZH

1.  W EH N * AO L * EH L S * F EY AX L Z * Y UX Z * F AO R S
2.  D OW N DX * AE S K * M IY * DX AX * K AE R IY * AX N * OY L IY * R AE G *
L AY Q K * DH AE Q T
3.  AX K R AO S * DH AX * S T R IY Q T * S T AE N Z * AX * K AH N T R IY * S K
UX L

1.  TH IH NG Z * IH N * AX * R OW * P R OW V AY DX * AX * S EH N S * AX V *
AO R D AXR
2.  AY L * M EY Q K * S EH N S * AX V * DH AX * P R AA B L AX M * IH N * AX *
M OW M AX N Q T
3.  DH AX * K OW Q T * HH AE Z * AX * B L EH N D * AX V * B OW TH * L AY Q
T * AX N * D AA R K * F AY B AXR Z

S OW
N AH CH
DH AE Q T
L AY Q T
DH EH R
HH AW S
S P EH SH AX L

1.  D UX * DH EY * G OW * AH P * AX N * D AW N
2.  Y UX * M AH S T * B L EH N D * S ER Q AX N * TH IH NG Z * T UX * M EY Q K *
AX
* S P EH SH AX L * W AE K S
3.  AX K R AO S * DH AX * S T R IY Q T * S T AE N Z * AX * K AH N T R IY * S K
UX L

1.  F AY V * S IH K S * EY Q * W AH N * N AY N * N AY N * EY Q T

2.  S IH K S * S EH V AX N * F AO R * F AO R * W AH N * S IH K S * S IH K S

WORDS 31
K AE SH
N AH TH IH NG
P OY N Q T
W AH Q T
S K UX L
DH AE Q T
B AX K AH Z

SENT 28
1.  DH AX * D AO R M AX T AO R IY * IH Z * B AX T W IY N * DH AX * HH AW S *
AX N * DH AX * S K UX L
2.  K AH M B AY N * AO L * DH IY * IH N G R IY D IY AX N T S * IH N * AX * L
AA R JH * B OW L
3.  D OW N Q T * AE S K * M IY * DX AX * K AE R IY * AX N * OY L IY * R AE G
* LL AY Q K * DH AE Q T

EMPH 1
P UH Q T * DH OW Z * T UX * B AE Q K
N OW * P UH Q T * DH IY Z * T UX * B AE Q K

WORDS 32
M EY JH AXR
F ER S T
AX B AW Q T
B AE Q K
DH IH S
N AH TH IH NG
W AE K S

SENT 29
1.  IH T S * JH AH S T * AX * L IH DX AX L * TH IH NG
2.  G R AE N M AH DH AXR * AW Q T G R UX * H ER * AH Q P B R IH NG IH NG *
IH N * P EH DX IY K OW Q T S
3.  D OW N Q T * AE S K * M IY * DX AX * K AE R IY * AX N * OY L IY * R AE G
* L AY K * DH AE Q T

HUNTER 77A
IH N * L EY Q T * F AO L * AX N * ER L IY * S P R IH NG * DH AX * SH AO R Q T
* R EY Z * AH V * DH AX * S AH N * K AO L * AX * T R UX * S AH N * AX V * DH
AX * AW Q T * AX V * D AO R Z * B AE Q K * T UX * DH AX * P L EY S AX Z * AH V *
HH IH Z * CH AY L DX H UH Q D */ T AA M * B R UH Q K S * W AH Z * S AH CH * AX
* M AE N * / IY CH * Y IY R * AE Q T * DH IY Z * T AY M Z * HH IH Z * D EH S K * M
EY D * HH IH M * W IH SH * F AO R * DH AX * L AY F * HH IY * N UX * AE Z * AX * B

OY */ IH N * DH AX * F AY V * Y IY R Z * S IH N S * L IY V IH NG * K AA L AX JH *
HH IY * HH AE D * N AA Q T * R IY V IH S IH DX AX D * HH IH Z * OW L Q D * HH AO
N Q T S * B IY F AO R */ B AH Q T * DH IH S * M AA R CH * T AA M * F AW N Q D *
HH IH M S EH L F * B AY * AX * S M AO L * S T R IY M * W IH TH * AX * G AH N

## HUNTER 77B

DH IH S * M AA R CH * T AA M * F AW N Q D * HH IH M S EH L F * B AY * AX *
S M AO L * S T R IY M * W IH TH * AX * G AH N * AE Q T * R EH S T * IH N * DH
AX * D R UH K * AH V * HH IH Z * AA R M */ DH AX * D EH S K * DH AE Q T * HH AE
Q D * T AY Q D * HH IH M * D AO N * W AH Z * G AO N * AX N * HH IH Z * W AH N *
TH AO Q T * W AH Z * F AO R * K W EY AX L */ HH IY * HH AE D * B EH N * AA N *
DH AX * T R EY AX L * S IH N S * D AO N * B AH Q T * N AA Q T * W AH N * B ER D *
HH AE D * K R AO S T * HH IH Z * P AE TH */ IH Q T * S IY M D * AE Z * DH OW * F
AY V * Y IH R Z * W IH TH AW Q T * HH AH N DX IH NG * HH AE D * M EY D * HH IH
M * L OW Z * T UH CH * W IH TH * AO L * DH AX * S M AO L * S AY N Z * DH AE Q T
* HH IY * W AH N S * N UX */ S AY N Z * DH AE Q T * W UH D * T EH L * F AO R * SH
ER * IH F * AX N * AE N IH M AX L * W AH Z * N IH R * AO R * N AA Q T

## HUNTER 77C

W AH N S * HH IY * TH AO Q T * HH IY * S AO * AX * B ER D * B AH Q T * IH Q T
* W AH Z * JH AH S Q T * AX * L AA R JH * L IY F * DH AE Q T * HH AE D * F AY
L Q D * T UX * D R AA Q P * T UX * DH AX * G R AW N D * D ER IH NG * DH AX
* W IH N DX AXR */ T AA M * S T AA P DX * N IH R * AX * S M AO L * S T R IY M * T
UX * R EH S T */ S UX N * AE F Q T AXR * HH IY * HH AE D * L EY DX * D AW N *
HH IH Z * G AH N * HH IY * HH ER D * DH AX * S AW N D * AX V * W IH NG Z * F R
AH M * AX K R AO S * DH AX * S T R IY M * AX N * F AY V * L AA R JH * D ER D Z *
K EY M * AW Q T * AX V * DH AX * B R AH SH */ DH EY * F L UX * T UX * DH AX *
EH JH * AX V * DH AX * S T R IY M * AH N AX W EH R * AX V * DH AX * HH AH N
DX AXR

## HUNTER 77D

DH AX * B ER D Z * F L UX * T UX * DH AX * EH JH * AX V * DH AX * S T R IY M *
AH N AX W EH R * AX V * DH AX * HH AH N DX AXR */ T AA M * P L EY S DX * HH
IH Z * HH AE N Q D * AA N * HH IH Z * G AH N * K W AY AX Q T L IY * S L OW L IY *
HH IY * R EY Z Q D * IH Q T * UX * HH IH Z * SH OW L D AXR * AX N * T UH K * EY
M */ DH AX * S EH K AX N D Z * T IH K D * AO F * L AY Q K * AW AXR Z * B AH Q T
* S T IH L * DH AX * B ER D Z * D R EY N Q K */ K W IH Q K * SH AA Q T S * R EY NG
* AW Q T */ DH AX * Y IH R Z * AX V * W EY DX IH NG * S IH M DX * T UX * D IH S
AX P IH R * W IH TH * DH AX * S AH Q K S EH S F AH L * K AH L M IH N EY SH AX N
* AX V * DH AX * H AH N Q T

## CLEAR

K AH M B AY N * AO L * DH IY * IH N G R IY D IY AX N Q T S * IH N * AX * L AA R
JH * B OW L

## WORDS 33

S IY M Q D
HH IH M S EH L F
P OY N Q T
HH EY AX L
W AE K S
D AO R M AX T AO R IY
HH EY AX L

SENT 30
1. Y UX * M AH S T * B L EH N D * S ER DX AX N * TH IH NG Z * T UX * M EY
Q K * AX * S P EH SH AX L * W AE K S
2. AY * TH IH NG K * DH AE Q T S * R IY L
3. K AH M B AY N * AO L * DH IY * IH N G R IY D IY AX N Q T S * IH N * AX * L
AA R JH * B OW L

SENT 31
1. P AO R K Y UX P AY N Z * R AX Z EH M B AX L * S IY * ER CH AX N Z
2. AH K R AO S * DH AX * S T R IY T * S T AE N Z * AX * K AH N T R IY * S K UX L
3. S EH K AX N * CH IH L D R AX N * AXR * AO F AX N * S P EH SH AX L *
P IY P AX L

SENT 32
1. P UH T *  DH IH S * W AH N * R AY T * HH IH R
2. CH IY P * S T AA K IH NG Z *  R AH N * DH AX * F ER S T *  T AY M * DH EY R * W
AO R N
3. DH AX * D AO R M AX T AO R IY * IH Z * B AX T W IY N * DH AX * HH AW S * AX
N * DH AX * S K UX L

WORDS 34
F ER S T
S EH K AX N D
HH EH D
L AO NG
D AO R M AX T AO R IY
DH AE Q T
R AY Q T

COUNT
W AH N * T UX * TH R IY * F AO R * F AY V * S IH K S * S EH V AX N * EY Q T * N AY
N * T EH N * AX L AX V EH N * T W EH L V * TH ER T IY N * F AO R T IY N * F IH F T
IY N * S IH K S T IY N * S EH V AX N T IY N * EY Q T IY N * N AY N T IY N * T W EH N
DX IY

WORDS 35
HH IH M S EH L F
K AE SH

P OY N Q T
K AA N V ER S EY SH AX N
DH AE Q T
SH UX Q T
HH AE Q D


<u>WORDS 36</u>
B AH Q T
Z IH R OW
M AH CH
P R OW G R AE M
HH IH M S EH L F
B OW TH
AX K R AO S


<u>WORDS 37</u>
HH EY AX L
L AY Q T
W AE K S
L AY Q T
SH IH P
B L EH N D
S K UX L


<u>NSEQS 5</u>
1.  F AO R * TH R IY * S EH V AX N * F AY V * W AH N * T UX * F AY V
2.  TH R IY * S IH K S * F AO R * S EH V AX N * N AY N * S IH K S * T UX
3.  W AH N * W AH N * F AO R * S IH K S * TH R IY * T UX * S EH V AX N


<u>FAST 3</u>
1.  DH AX * AH DH AXR * W AH N * IH Z * T UX * B IH G
2.  K AX M B AY N * AO L * DH AX * IH N G R IY D IY EH N T S * IH N * AX * L AA R
JH
* B OW L


<u>EMPH 2</u>
1.  P UH Q T * DH IY Z * T UX * D AW N
2.  N OW * P UH Q T * DH IY Z * T UX * B AE Q K


<u>WORDS 38</u>
HH IH M S EH L F
B L IH NG K
B AH Q T
AX B AW Q T
M EH ZH ER
K OW Q T

W AE K S

SENT 33
1.  D AH Z * K R IY OW L * K UH K IH NG * Y UX Z * K ER IY
2.  AX * R OW L * AX V * W AY AXR * L EY * N IH R * DH AX * W AO L
3.  D OW N Q T * AE S K * M IY * DX AX * K AE R IY * AX N * OY L IY * R AE G
* L AY Q K * DH AE Q T

SENT 34
1.  DH IY * AH DH AX * W AH N * IH Z * T UX * B IH G
2.  IH F * AY * HH AE D * DH AE T * M AH CH * K AE SH * AY D * B AY * DH A *
HH AW S
3.  Y UX * K AX N * SH UX T * AE T * DH AX * SH IH P * AXR * D UX * N AH TH IH
NG

SENT 35
1.  IH F * AY * HH AE D * DH AE Q T * M AH CH * K AE SH * AY D * B AY * DH AX *
HH AW S
2.  DH IY * HH AE D * Y ER * D AA R K * S UX DX * AX N * G R IY S IY * W AA SH *
W AA DX AXR * AO L * Y IH R
3.  P UH Q T * DH IY Z * T UX * B AE Q K

WORDS 39
W AE K S
AO R D AXR
K OW Q T
R AY Q T
P R AA B L AX M
F ER S T
M AH CH

WORDS 40
HH IH M S EH L F
B AE Q K
S T R IY T
N AH TH IH NG
B AE Q K
TH IH NG Z
S T R IY T

SENT 36
1.  EH L D ER L IY * P IY P AX L * AA R * AO F AX N * AX K S K L UX D AX D
2.  W EH N * AO L * EH L S * F EY AX L Z * Y UX Z * F AO R S
3.  DH AX * D AO R M AX T AO R IY * IH Z * B AX T W IY N * DH AX * HH AW S * AX
N * DH AX * S K UX L

LISTENING ASSESSMENT: JW54

ERROR CODES:

| | | | | | |
|---|---|---|---|---|---|
| [01] | = | phoneme deletion | [11] | = | sound-syllable repetition |
| [02] | = | phoneme distortion | [12] | = | sound-syllable revision |
| [03] | = | phoneme substitution | [13] | = | word repetition |
| [04] | = | word substitution | [14] | = | word revision |
| [05] | = | word addition | [15] | = | phrase repetition |
| [06] | = | word deletion | [16] | = | phrase revision |
| [07] | = | word transposition | [17] | = | truncated |
| [08] | = | inappropriate prosody | [18] | = | dysfluent reading |
| [09] | = | hyperarticulation | [19] | = | phoneme addition |
| [10] | = | inappropriate pause | [20] | = | phrase addition |

words0l

    problem[03;pl/pr]
    school[17]

words02

    children[02;/ch/]

words04

    quite[17]

sent02

"2. Don't do Charlie's[02;/ch/,/s/] dirty dishes."
"3. She had your dark suit[02;/s/] in greasy wash water all year."

gfathr#01 la

"You wish to know all about my"
"grandfather.  Well, he is nearly 93"
".years old, yet he still[02;/s/] thinks[02;/s/I"
"as[02;/s/I swiftly as ever.  He dresses himself in"
"an old black frock coat, usually several"
"buttons missing.  A long beard clings[02;/s/I to"
"his chin, giving those who observe him a"
"pronounced[02;/s/I feeling of the utmost[02;/s/]"

"respect[02;/sp/l.  When he speaks, his voice"
"quivers a bit.  Twice each day[17] he plays"
"skillfully upon a small organ."

gfathr#0l lb

 "Twice each day he plays skillfully and"
"with zest upon a small organ.  Except in"
"the winter when the snow or ice prevents,"
"he slowly takes a short walk in the open"
"air each day.  We have often urged him to"
"walk more and smoke less, but he always"
"answers, 'Banana oil!' Grandfather likes"
"to be modern in his language."[02;all /s/ and
/s/ blends]

vowels#013

ay (date)        ah (hot)[03;aw/ah]              ae (bat)

vcv's#015

  uhfa          uhzA[12; "uhr..uhza"]          uhwa

sent03#016

"1. Don't ask me to carry an oily[ll;"oi..oily"] rag like that."
"2. You can shoot[02;/sh/I at the ship[02;/sh/]
  or do nothing."
"3. You must blend certain[13,02;/s/I things[02;/s/I
  to make a special wax."

words08

        words[02;/s/]
        first[02;/s/]

sent04

"2. Across the street stands a country "
   school."[02;on all /s/]
"3. The dormitory is[02;/s/] between the
  house[02;/s/] and the school."


sent05

" 1. I assume moisture will damage this[02;/s/] "
   ship's[02;/sh/] hull.  "
"3. If I had that much[02;/ch/I cash I'd buy the
   house.  "

words10

        school[02;/sk/]
        seemed[02;/s/]
        house[02;/s/]

sent06

"l. Things in a row provide a sense[02;/s/] of "
   order.  "
"3. Second[02;/s/] children are[ll;llar..arell]
  often[IO] special people."

words 12

        children[02;/ch/]

slowl

" 1. The other one is[02;/s/ on both reps] too big."
"2. Combine all the ingredients[02;/s/ on both reps] in a
  [12;"s..Iarge"] large bowl."

wordsl3

        seemed[02;/s/]
        yourself[02;/s/]
        across[02;/s/I
        sense[02;/s/]
        second[02;/s/I

wordsl4

        this[02;/s/]
        across[02;/s/I

sent07

"l. When all else[02;/s/I fails[02;/s/], use force[02;/s/]."

125

sent09

"2. I'll make sense[02;/s/] of the problem in a "
   moment.  "

wordsl5

      sense[02;/s/]
      people[17]

wordsl6

       beautiful[03;i/1]
      sense[02;/s/]
      second[17,02;/s/]

fast 1

" 1. Combine all the ingredients in a
        large[02;/j/,both reps] bowl."
"2. The[10;on 2nd rep] other one is too big."

wordsl7

      smooth[02;/sm/1
      special[02;/sp/]

slow2

" 1. The other one is[02;/s/,on 2nd rep] too big."
"2. Combine all the ingredients[02;/s/,on 2nd rep] in a
  large[17; on 2nd rep] bowl."

sent l 1

"1. When all else[02;/s/] fails[02;/S/], use force.[02;/s/]"
"2. Combine all the ingredients[02;/s/] in a "
   large bowl."
"3. Things[02;/s/I in a row provide a sense[02;/s/] of "
   order.  "


sent 1 2

" 1.  You must blend certain[02;/s/] things[02;/s/] to make "
     a special[02;/sp/] wax.  "

126

"2.  The dormitory is between the house[02;/s/] "
      and the school." [02;/s/I
"3.  Don't ask[02;/sk/I me to carry an oily[ll;"oi..oily"] rag " like that.  "

wordsl9

        ship[02;/sh/I
        house[02;/s/I
        things[02;/s/I

fast2

"l. Combine all the ingredients[02;/s/,on both reps]
  in a large[02;/j/,on both reps] bowl."

 "2. The other one is[02;/s/,on both reps] too big."

sent 1 3

" 1. You can shoot[02;/sh/I at the ship[02;/sh/I or do "
  nothing.  "
"2. The gorgeous[02;/s/I butterfly ate a lot of "
    nectar.  "
"3. You must blend certain[02;/s/] things[Ol;/s/,10] to make
    a special[02;/sp/] wax."

words20

        seemed[02;/s/I
        become[ll;llbe..become"I
        seemed[02;/s/I

sentl4

" 1. The other one is[02;/s/] too big."
"2. She[02;/sh/] always jokes[02;/s/I about too much[02;/ch/] garlic in his food."
"3. If I had that much[02;/ch/] cash[02;/sh/I
        I'd buy the house.[02;/s/]"

sentl5

"2. Across[02;/s/I the street[02;/st/] stands[02;/st/]
  a country school.  " [02; /s/I
"3. She had your dark suit[02;/s/] in greasy[02;/s/]

127

wash water all year."

words2l

      second[02;/s/]
      street[02;/s/]

sent 1 6

"2. Second[02;/s/] children[02;/ch/] are often
  special [02; /s/I people."
"3. I'll make sense[02;/s/] of the problem in a
   moment.  "

sentl7

"3. Things[02;/s/] in a row provide a sense[02;/s/I of "
   order.  "

 nseqs2

   7789388 8761335[02;/s/ in 17' and '6'] 2918524

sentl9

"3. The coat has[Ol;/s/] a blend of both light "

   and dark fibers."

sent20

" 1. The sermon[02;/s/] emphasized[02;/s/] the need for "
   affirmative action."
"2. Second[02;/s/] children[02;/ch/I are often
 special[02;/s/I people.  "
"3. That noise problem grows more "
   annoying each[02;/ch/] day.  "

slow3

"l. Combine all the ingredients[02;/s/I in a
   large[02;/j/1 bowl.  "
"2. The other[15;"the other..the other"] one
 is[02;/s/I too big."

sent2l

" 1.  The coat has[02;/s/I a blend of both light " and dark fibers[02;/s/].  "
"2.  They all know what I said[02;/s/]."
"3.  I'll make sense[02;/s/I of the problem in a " moment.  "
"4.  Combine all the ingredients[02;/s/] in a " large[02;/j/1 bowl.  "

sent22

1.   Things[02;/s/I in a row provide a sense[02;/s/I of "
          order.  "
" 2. Put these two back."
"3. She had your dark suit[02;/s/I in greasy[02;/s/] wash
    water all year."



words26

        much[02;/ch/]

 words27

        first[02;/st/]

nseqs3

   6582269[02;/s/ in '6'] 7217[10]424 2315483

sent23

"3. Second[02;/s/I children[02;/ch/j are often

        special[02;/s/j people."

sent24

" 1. When all else[02;/s/] fails[02;/S/I, use force[02;/s/].  "
"2. They all know what I said[02;/s/]."
" 3. The oasis[02;/s/] was a mirage."

sent25

"3. Across[02;/s/I the street[02;/s/I stands a country "
   school. [02;/s/I "

sent26

"l.   Things[02;/s/] in a row provide a sense[02;/s/] of " order.  "
"2.   I'll make sense[02;/s/] of the problem in a " moment.  "
"3.   The coat has[Ol;/s/I a blend of both light "
      and dark fibers."

sent27

"3. Across the street[02;/s/] stands[02;/s/j a country "
    school[02;/s/l.  "

sent28

"2. Combine all the ingredients[02;/s/] in a
    large[02;/j/] bowl."

emphl

"Put those two back?"
"No, put THESE[16; "put tho..put these"] two back."

 words32

        first[02;/s/]

sent29

" 1. It's[Ol;/s/] just[02;/j/] a little thing."
"3.  If I had that much[02;/ch/] cash I'd buy the
     house.  "

hunter#077a

"In late fall and early spring the short"
"rays of the sun call a true son of the"
"out-of-doors back to the places of his"
'.childhood. Tom Brooks was such a man."
"Each[02;/ch/] year at these times his desk seemed"
"like a stone whose weight made him wish"
"for the life he knew as a boy.  In the"
"five years since leaving college he had"
'.not revisited his old haunts before."
"But this March Tom found himself by a[17]"
"small stream with a gun."[02;on all /s/ and /sh/ throughout passage]

hunter#077b

130

"This March, Tom found himself by a"
"small stream with a gun at rest in the"
"crook of his arm.  The desk that had tied"[10]
"him down was gone and his one thought"
"was for quail.  He had been on the trail"
"since dawn, but not one bird had crossed"
"his path.  It seemed as though[10] five years"
"without hunting had made him lose touch"
"with all the small signs that he once"
"knew - signs that would tell for sure if[17]"
"an animal was near or not." [02;on all /s/ and /sh/ throughout passage]

hunter#077c

"Once he thought he saw a bird, but it"
"was just a large leaf that had failed"
"to drop to the ground during the winter."
"Tom stopped near a small stream to"
"rest.  Soon after he had laid down his"
"gun, he heard the sound of wings from"
"across the stream, and five large birds"
"came out of the brush.  They flew to the"
"edge of the stream unaware of the"
"hunter." [02;on all /s/ and /sh/ throughout passage]


hunter#077d

"The birds flew to the edge of the stream"
"unaware of the hunter.  Tom placed his"
"hand on his gun quietly.  Slowly he"
"raised it to his shoulder and took aim."
"The seconds ticked off like hours, but"
"still the birds drank.  Quick shots rang"
"out.  The years of waiting seemed to"
"disappear with the successful"
"culmination of the hunt." [02;on all /s/ and /sh/ throughout passage]

clear#078

"Combine all[05;llofll,on 3rd rep] the ingredients[02;/s/,on all reps]
in a large[02;/j/,on all reps] bowl.  "[17;only 4 reps]

sent30

"3. Combine all the ingredients[02;/s/] in a

large[02;/j/] bowl.  "

sent31

" 1. Porcupines resemble[02;/S/l sea[02;/s/] urchins[02;/sh/].  "
"2. Across[02;/s/] the street[02;/s/] stands[02;/s/I a country "
   school. [02; /s/I "
"3. Second[02;/s/I children[02;/ch/I are often
  [15;"are of..are often"] special people."

sent32

" 1. Put this[02;/s/I one right here."

words34

        first[02;/st/]

words35

        himself[02;/s/I

words36

        himself[02;/s/I

 nseqs5

        4375125 3647962 1146[101327

fast3

" 1. The other one is[02;/s/] too big."
"2. Combine all the ingredients[02;/s/] in a
   large[02;/j/] bowl."



emph2

"Put these two down?"[17;on 5th rep]
"No, put these two BACK."

sent34

" 1. The other one is[02;/s/] too big."

"2. If I had that much[02;/ch/] cash[02;/sh/]
  I'd buy the house."
"3. You can shoot[02;/sh/] at the ship[02;/sh/j
  or do nothing."

sent35

" 1. If I had that much[02;/ch/I cash[02;/sh/]
  I'd buy the house."
"2. She had your dark suit[02;/s/] in greasy wash
    water all year."

words39

        first[02;/st/]

sa#101

        sa, sa, sa..."[02;/s/]

**APPENDIX E**:  Mistrack log

Information regarding the JW mistrack files:

Mistrack files contain the ID number of the task, a pellet identifier, and the onset and offset times of the mistracks, in milliseconds from the start of the record.  Mistracks can be categorized as either NOT FOUNDS or RASTER HOPS.  NOT FOUNDS occur when the microbeam looks for a pellet but is not able to find it.  RASTER HOPS occur when a raster follows the wrong pellet: e.g., when the T1 raster follows the T2 pellet.  To save keystrokes we devised a code for RASTER HOPS using the letters 'a' thru 'k'. Each pellet is assigned a letter in the following manner.

    UL   (a)
    LL   (b)
    MaxI (c)
    MaxN (d)
    MaxG (e)
    ManI (f)
    ManM (g)
    T1   (h)
    T2  (i)
    T3  (j)
    T4   (k)

The RASTER HOPS are indicated in the column labeled "Raster."  For example, if the T2 raster followed the T3 pellet, the code would be i;j.  That is, "T2 raster goes to the T3 pellet."  In this case, both x and y time histories for channels labeled T2 and T3 would have the exact same data in them (T3 positions), and the T2 data would be lost.   Sometimes however, another raster may happen to hop to the "lost" pellet (T2 in this case) and track it.  For instance, the T1 raster may hop to the T2 pellet.  In this case, the data in the time series labeled T1 will be movement of the T2 pellet.  We have tried to switch channel names to account for this situation, but at present this has not been consistently done.  Occurrences of switches like this, where good data has a wrong channel name, are coded in the "raster" column.  For the example of the RASTER HOP between the T1 raster and the T2 pellet, the coding in the "raster" column would be h;js.  The 's' stands for a channel name switch between the two channels. The pellet movement for T2 (letter 'j') actually resides in the channel labeled T1 (letter 'h').

Mistrack File: JW54:

| Rec ID Number | Pellet | Onset | Offset | Raster |
|---|---|---|---|---|
| 32 | T3 | 2007 | 2358 | |
| 32 | T1 | 0 | -1 | |
| 35 | T2 | 4915 | 4960 | |
| 35 | T1 | 2580 | 2820 | |
| 37 | T4 | 0 | 3650 | |

| | | | | |
|---|---|---|---|---|
| 37 | T1 | 0 | 1840 | |
| 41 | T3 | 1980 | 2170 | |
| 41 | T4 | 19433 | 19681 | |
| 49 | T2 | 9000 | -1 | i;g |
| 53 | T1 | 0 | 130 | |
| 55 | T2 | 3050 | -1 | i;g |
| 59 | T1 | 0 | 2075 | |
| 61 | T4 | 2890 | 3160 | |
| 64 | T2 | 3750 | 3810 | |
| 72 | T2 | 6615 | -1 | i;g |
| 73 | T2 | 732 | 781 | |
| 80 | T3 | 7240 | 7450 | |
| 82 | T2 | 3500 | 5950 | |
| 100 | UL | 2035 | 2135 | |
| 103 | UL | 635 | 725 | |
| 104 | T4 | 0 | 2185 | |
| 108 | T4 | 0 | 18950 | |
| 109 | T4 | 5810 | 6235 | |
| 109 | UL | 946 | 1028 | |
| 110 | T4 | 15790 | 16000 | |
| 113 | T3 | 5500 | 5675 | |
| 113 | T4 | 0 | 5600 | |
| 116 | T2 | 1800 | 1980 | |
| 116 | T2 | 1981 | -1 | is;h |
| 116 | T1 | 0 | -1 | h;c |
| 120 | UL | 586 | 675 | |
| 121 | UL | 2860 | 3425 | |
| 123 | T3 | 4395 | 4450 | |
| 127 | T4 | 0 | 1440 | |
| 129 | T4 | 0 | 5640 | |
| 131 | LL | 1866 | 3973 | |
| 138 | MANI | 2581 | 2675 | |
| 138 | MANI | 3736 | 3812 | |
| 138 | MANI | 4855 | 4967 | |
| 141 | T4 | 2025 | 2115 | |
| 149 | T1 | 0 | 1720 | |
| 154 | T2 | 0 | 540 | |
| 154 | T2 | 541 | -1 | i;j |