

VOICE RECOGNITION BY MAN, ANIMAL, AND MACHINE

HENRY M. TRUBY

This report treats my own personal interpretation of the implications of Voice Recognition... for Speech and for all analogous intraspecies and interspecies communication, for Diagnosis whether or not related to speech, and for Law in a restricted but universally significant sense. Having protected myself with the above specifications, I can safely generalize by saying that this paper offers my own interpretation of the implications of Voice Recognition for Communication, Diagnosis, and Law, in terms of its scope, its artifacts, and its prospects, as I presently see them.

I have stressed above, and will reiterate, the 'personal' aspect for two reasons. First, I make no attempt here to review whatever 'voice recognition literature' there might be, nor to criticize the hypotheses or efforts of others in any relevant regards; and second, it has been made very clear that on certain of the issues relevant to Voice Recognition in the immediately passed years, I have found myself to be standing essentially alone among my contemporaries — resolutely and outspokenly, but nonetheless alone, as will be elaborated below. The chief controversial issues, notably, are those of infant-cry-sound analysis or "Cryprinting" (Truby 1960, Truby, Bosma, Lind, and Karlberg 1960, Truby, Lind, and Bosma 1961, Truby 1962a, 1965c, Truby, Bosma, and Lind 1965, Truby 1966b, 1967a, 1967d, 1967g, 1967h, 1970a), of so-called 'speaker identification' or 'Voiceprinting' (Kersta 1962, 1969, Kersta and Colangelo 1969, Truby 1970a), and of "'man-to-dolphin and dolphin-to-man' translation" (Lilly, Truby, *et al.* 1967, Truby 1967b, 1967c, 1970b, Truby and Lilly 1967). I see no need as yet for NAMING the specific area treating the *acoustigraphic analysis of the sounds of other-than-human creatures and am content to relate all such visible-acoustic data to the broad area I shall refer to, at least for the time being, as 'soundprinting', to include 'cryprinting' and 'voiceprinting'.

The second reason that I stress the 'personal' aspect is in order to illustrate the network of involvements and associations over the past 25 years which has inspired me or contrived to force me to optimism which few of my colleagues, known or unknown to me, seem ready to accept or share.¹

¹ With a backlog of twenty-five years of university teaching and research and over a hundred publications, I have at least observed how really personal is most scientific reporting after all. Too

The above-mentioned rubrics implicated by Voice Recognition, namely Communication, Diagnosis, and Law, are not a trichotomy — i.e., they are not mutually exclusive. As suggested in my opening sentence, COMMUNICATION intends all forms of animate sound-making and -receiving, including speech and *Pseudospeech, which last is my term for the so-called ‘synthetic speech’ of the communications engineer or the ‘simulated speech’ of the computer (I’m certain that every phonetician would agree that the output of so-called ‘synthesizers’ is not SPEECH — nor even ‘synthesized speech’ — but rather *PSEUDOSPEECH).² DIAGNOSIS is presently almost entirely — but not exclusively — limited to HUMAN considerations and implicates not only the detection and evaluation of so-called “speech disorder” (Truby 1966a), “voice disorder” (Truby 1970a), and the like, but also — and, I believe, more importantly — the detection and evaluation of physiological and pathological state (e.g., Karelitz, Karelitz, and Rosenfeld 1959, Truby *et al.* 1960, Truby, Lind, and Bosma 1961, Karelitz and Fisichelli 1962), this last from such forms of vocal output as neonatal and subsequent infant-cry-sound (Truby 1960, Truby, Bosma, Lind, and Karlberg 1960, Truby, Bosma, and Lind 1965a, Truby, Lind, and Bosma 1961) from all other prespeech and non-speech phenomena, and from all PARA-LINGUISTIC vocal output; and LAW implicates that sector of interhuman affairs which is concerned not only with the establishment or corroboration of individual identity (Bolt *et al.* 1969, Kersta 1962, 1969, Kersta and Colangelo 1969, Ladefoged and Vanderslice 1967, Tosi *et al.* 1971, Truby 1970a, and Stevens’ contribution to this Congress, to be found on pp. 206-232 of this volume) but as well with the validity of witness testimony — i.e., exclusively human, *sans* instruments, witness testimony — regarding identifications made on the basis of voice characteristics and/or speech characteristics. Thus, the ‘witness testimony’ to which I refer is that of ‘the ordinary citizen’ — NOT ‘the expert witness’.

Often the revelations of predecessors are ignored, as was demonstrated so succinctly in Panconcelli-Calzia’s marvelous little book (1941) and “each new star on the horizon must rediscover for itself all the time-worn reaches of its own flickering light” (Truby 1970a). Especially re instrumental speech-sound analysis, a careful trip through the latter half of the 19th century with Wheatstone, Grassman, Willis, Helmholtz, and perhaps especially, Scripture, would lighten much of the load presently carried by experimental phoneticians of WHICHEVER ilk and pedigree, and in the same idiom of ‘phonetics via physics’, I recommend the excellent 1940 book and 1924-7 series of “professor of vocal physiology” Mark H. Liddell. The Bulletins, interestingly enough, were published by the Purdue University Engineering Experiment Station, and note the title, especially the last word!: *The Physical Characteristics of Speech Sound* (Lidell 1924-7, 1940).

² As is discussed in detail in Truby 1967e: 520, 538-9, 551-2, “synthetic speech” depends for its effectiveness upon its *caricature* nature (Truby 1965a) and is more aptly termed PSEUDOSPEECH (see also Truby 1964c, 1965b, 1965d, in press, 1967e, and 1967f), the implications of which should be immediately evident, at least to linguists. Machines simply do not produce SPEECH, and ‘analysis by synthesis’ procedures are precariously optimistic where based on the assumption that the responses of supposedly ASSORTED — but invariably linguistically and acuitously DIVERSE — human listeners to ‘synthetic-speech’ output are sufficiently analogous to corresponding responses to HUMAN speech as to infer validity in the artifact. Valid terms: ‘synthetic speech’ and ‘simulated speech’ (if unabridged) and ‘pseudospeech’. Dangerous or invalid terms: ‘synthesized speech’, ‘speech synthesis’, ‘speech synthesizer’, ‘speech simulation’, etc. (cf. Dudley 1936 and 1939). And see Truby 1971c.

I hope that my Abstract appearing in the Program of this Congress seems neither trite on the one hand nor enigmatic on the other, and I shall attempt to elaborate its particulars below, first by here repeating it in its entirety:

Man has been universally aware of the individuality of individual voice performance since at least the dawn of written history. Intraspecies and interspecies recognition by animals of species specific and human ‘voice’ respectively has been demonstrated generally through history and instrumentally since the advent of phonographic recording. And transmission monitoring instrumentation is presently adequate, in the proper hands, to the specification of idiosyncratic particulars demonstrably present in the speech signal, however notably subliminal to the still frustrated ‘phoneme detector’.

Experimental phoneticians, as well as scientists of other disciplines, have seriously and actively pursued a voice recognition procedure for at least the past century. The author and other basic researchers with the sound spectrograph began focusing redoubled attention on this aspect of idiolectalism some twenty-five years ago. During the most of this period the author particularly has reinforced his own considerations of voice idiosyncrasy via ‘speech synthesis’ variously, paranatal and neonatal cry-sound analysis, and sonocine-radiography of speech and of cry as noted, with an instructive excursus in an attempted analysis of the complex communication network of the Bottlenose Dolphin.

There is nothing more subjectively personal than the interpretation — to include ‘diagnosis’ — by one individual of another individual’s performance, vocal or other; and one man’s opinion is clearly NOT as good as another’s in pattern recognition affairs. The report offered this Congress is primarily an autobiographical account of the author’s concern with voice identification, rather than an attempt to appraise the general scientific attitude regarding any specific instrumentation or interpretation such as ‘voiceprinting’, an aspect of voice analysis which — like ‘cryprinting’ — the author has steadfastly maintained has explicitly definable validity...again, ‘in the proper hands’. Certainly, the identificatory reliability attributed voice identifying ‘witnesses’ in legal procedures warrants careful reconsideration. The visible-acoustic parameters functioning in human voice manifestation on sound spectrograms were set out by the author in 1959 in terms of nine directly measurable DESCRIPTIVE FEATURES in a book which includes 1100 sound spectrograms (correlated with 4400 other acoustigrams). The resultant parametric sets, or PHONETTES, indicate a physical framework for the scientific application of the sound spectrogram to individualistic voice recognition by man — and, in extension, by machine...the relevant computer program being but a hypothetical step away.

I think of the three paragraphs of my Abstract as constituting a prelude, an introduction, and a discussion. To these I have appended some remarks which hopefully will be accepted as a conclusion.

There is much to be mentioned all-at-once, in these regards, and I shall certainly try to at least touch upon as many as possible of these details of equal rank, with the limitations always imposed by the TEMPORAL LINEARITY of speech.

Before reverting explicitly to the elements of my Abstract, I feel I must interpose a few words about VOICE. For instance, taken in its broadest sense — or attributing such a broad sense to it, ‘voice’ has been applied not only to the speech signal generated in *Homo sapiens* but also to quasi-analogous audible-signal generation in other animate — as opposed to plant — species. Thus it is that among the members of the animal kingdom, men, almost all so-called ‘lower animals’, MOST birds, and SOME

INSECTS are said to have 'voices' characteristic of the relevant particular species. Beyond that, within each species there is apparently individuation, so that at least hypothetically, no two — even species-specific — 'voices' are identical, and each individual 'voice' appears to be individualistic... unique... idiosyncratic, as far as we can tell to date, that is, insofar as is observable in man's current archives of such physically measurable output.

When I say 'as far as we can tell', I mean also that our present sound-analysis instrumentation has been developed to the point of indicating physical evidence of discrimination among the 'voices' of individuals, I believe, even though as yet, COMPLETE analyses of these sets of differences have not been made, to my knowledge.³

Certainly my own efforts over the past twenty-five years at defining human speech were begun with at least the awareness implied above, thanks in large part at the beginning, to the exceptional tutelage of the late Miles L. Hanley, phonetician extraordinary and charter member of these Congresses and of the International Phonetic Association.⁴ In any event, I personally was already definitely convinced, in 1945 (with the public announcement and appearance of the Bell Laboratories sound spectrograph, that is [Potter 1945]), that the individual voice has an individuality to be depended upon... and is ultimately qualifiable. And what's more, I felt sure, already at that time, that this individuality could be demonstrated with optimal sound-spectrographic analysis and appropriate interpretation. Please, hang on the word 'optimal', and remember that this instrumentation indicated directions in which acoustic analysis seemed destined to go.

In 1946, I was an instructor of engineering mathematics at the University of Wisconsin when Hanley put into my hands accounts of the Fourier analysis of speech-sound AND the memorable 1934 article by John Steinberg in the *Journal of the Acoustical Society of America*, describing the drawing-board forerunner of the sound spectrograph at Bell Telephone Laboratories. I can remind you, in retrospect, that 2 SECONDS of speech-sound continuum required from two to as much as ten HOURS of manual measurement and conversion — a feat accomplishable in ONE MINUTE with a sound spectrograph! Having ACCOMPLISHED something in one minute, however, rarely means that interpretation of the relevant details has been speeded-up, or even necessarily improved.

Now, Miles Hanley and I spent many long hours in 1946 scrutinizing the vast

³ Now, 'voice', as used far too generally, is innocently identified in humans with all the signal features more accurately specified as SPEECH features (dialectal and/or idiolectal) and in non-human animals as such pattern-specific features as so-called 'birdcalls', dolphin 'whistles', wolf and other 'howls', 'roars', 'rattles', 'whinnies', 'growls', 'snarls', 'screeches', 'hisses', and so forth, anthropomorphically speaking. However specified, these are certainly the bases for intra- and inter-species identification. BUT, each individual has 'its' own intimate set of familiar receivers, each of whom is able to recognize THE PARTICULAR SENDER, with varying limitations. It is THIS parameter of individualism for which the term 'voice' has relevance, as will be enlarged upon later in this report.

⁴ Also Associate Director, *Linguistic Atlas of New England* (1939), lexicographer, and dialectologist, for which see Truby 1967e (footnote 13) and biographical directories.

Gibbs and Gibbs *Atlas of Encephalography* (Gibbs and Gibbs 1944, Truby 1967e), in our search for physiological correlates of linguistic and paralinguistic features in the EEG, before turning to sound spectrograms. These were very exciting times for acoustic analysis: Martin Joos was also at Wisconsin and working on his incomparable monograph *Acoustic Phonetics* (Joos 1948); Armando de Lacerda was a visitor there — Lacerda had invented his ingenious Polychromograph (Lacerda 1933, 1934; Truby, Bosma, Lind, and Karlberg 1960, Truby, Lind, and Bosma 1961, Truby, Bosma, and Lind, 1965, footnote 23) forerunner of the ink-jet oscillograph, some 15 years previously; and it was also in 1946 that the comprehensive Bell Labs monograph *Technical Aspects of Visible Speech* appeared (Dudley and Gruenz 1946, Koenig, Dunn, and Lacy 1946, Kopp and Green 1946, Potter 1946, Riesz and Schott 1946, Steinberg and French 1946; Potter, Kopp, Green 1947; Truby 1959, 1962b and 1967e, fn. 14). This monograph to this day and almost without exception constitutes a complete presentation of sound-spectrographic analysis — its implications, its applications, its technical definition, and its physiological reference — and is in addition a simplified yet thorough description of sound-spectrographic analysis and the sound spectrograph (Koenig, Dunn, and Lacy 1946), and a clarification of the significance of feature details appearing on sound spectrograms — both as to acoustic and to phonetic correlates, plus an entire classified gallery of the visible correlates of representative 'individual speech sounds' (Kopp and Green 1946) of so-called 'American English', and an illustrative display of the appearance of many of these and other visible patterns as they occur in connected speech (Truby 1962b). One doesn't need to do all that again. That has been done.

In 1947 the second book in our history to be called *Visible Speech* appeared. The FIRST *Visible Speech* was published in 1867 and is a 150-page treatment of "The Science of Universal Alphabets", by Alexander Melville Bell (1867).⁵ The 1947 *Visible Speech* is by Ralph Potter, who originally proposed the sound spectrograph, George Kopp, who was the only person I've known who claimed to be able to 'read' sound spectrograms, and Harriet Green, who later became Mrs. Kopp (Potter, Kopp, and Green 1947). In this book we find an organized, step-by-step discussion, without technical embellishment, of PHONES... in terms of their corresponding physical characteristics as seen on sound spectrograms (Truby 1962b). In all of my discussion this morning my concern is with PHONES... and only by implication with 'phonemes', depending on how you interpret them. I hope to illustrate the dichotomy below. In the 1948 monograph by Joos we find not only expert knowledge of the subject of sound-spectrographic analysis and a thorough coverage of technical details, but also a selection of hints on the practical problem of making, interpreting, and otherwise dealing with sound spectrograms (Truby 1962b). With the advent of the first commercial sound spectrograph, the Kay Sona-Graph, circa 1949, I, for one,

⁵ The alphabetic systems of Bell, Jespersen, Lundell, and others are of dialectological — even idiolectological — application, and thus of relevance to 'speaker identification', but they do NOT implicate the VOICE.

began to note spectral characteristics on Sonagrams which I had not noticed on the previous Bell Labs sound spectrograms, cut-off as they were below 200 Hz and above 3600 Hz, and it became obvious to me that certain idiosyncracies of vocal-tract performance were apparent from individual to individual speaker.

By the end of 1950, in Harvey Fletcher's laboratory at Columbia University, I had made some 3,000 sound spectrograms of my own speech in a preliminary effort at understanding the implications and complexities of linguistic analysis in the sound domain and at clarifying the nature of the aforementioned speaker idiosyncracies.

These individualities were even more dramatically revealed for me at Haskins Laboratories, from 1950 to 1954, when in designing PB-2 ('Pattern Playback') compensatory-pattern constructs on the basis of Haskins sound spectrograms, it was clear to me WHOSE SPEECH was the source of the patterns, when real speech was mimicked in the synthesizing process — and, when synthesized 'from scratch', whether it was Pierre Delattre or I who had actually done the synthesizing! That is, even in the SYNTHETIC VERSIONS — and thus one dimension removed, reflections of speaker individuality were apparent! These observations were next corroborated — and extended to voice itself — on my very first Ove-2 pattern constructs in Stockholm in 1955 through 1957, where similarly the patterns synthesized on the basis of MY voice or of Gunnar Fant's voice disclosed the particular source *voice* involved... even when transmitted over the laboratory intercom! Or I might say ESPECIALLY when transmitted over such low-fidelity systems... not to exclude the *telephone*...

In 1957 in Stockholm, I launched the newborn-infant cry-sound studies,⁶ and in 1958 these and much preoccupation with segmentation (Truby 1958a, 1958b, 1958c, 1959, 1963, 1964a, Truby and Wegelius 1958), with publishing a book plus sound-spectrogram gallery (Truby 1959), and with the preparation of Fant's doctoral dissertation contrived to keep 'speaker recognition', etc., considerations on my list of things to be looked-at again and further *someday*.

The most comprehensive, visible-acoustic description that I know of, of at least ONE mode of vocal performance, is the initial definitive article in the pediatric field on Infant Cry Sound — an article growing out of my own research in Stockholm, with Drs James Bosma and John Lind (as elaborated in footnote 6). First published in Stockholm in 1960 as "Infant Cry Sounds: A Visual-Acoustic Analysis Technique"

⁶ Specifically, the sound-spectrographic analysis of paranatal and neopostnatal cry-sound, recorded via direct larynx auscultation, to identify brain damage AT BIRTH or as soon as possible thereafter (Truby, Bosma, and Lind 1965, Truby, Bosma, Lind, and Karlberg 1960, Truby, Lind, and Bosma 1961, Truby 1960, 1965c, 1966b, 1967d, 1967g, 1967h, Bosma, Truby, and Lind 1962, 1963), in association with Pediatricians John Lind, M.D. and James Bosma, M.D., at The Wenner-Gren Cardiovascular Research Laboratory, under sponsorship of The Swedish Medical Research Council, the Association for the Aid of Crippled Children, the United Cerebral Palsy Research and Educational Foundation, and the National Institute of Neurological Diseases and Blindness. The cryprinting aspect of this pediatric venture was not only the first application of 'voiceprinting' — in its most technically specific sense — but constitutes a vital corroboration of voice individuation (and see Gray and Kopp 1944).

(Truby, Bosma, Lind, and Karlberg 1960, plus tape) and retitled and submitted as "Cry Sounds of the Newborn Infant" in 1961 (Truby, Lind, and Bosma 1961, and see Truby, Bosma, and Lind 1965), this 59-page report treats our three years (at that time) of birth studies IN Stockholm, and comprises 80 cry-sound spectrograms and correlated textual descriptions, and a 19-minute, phonograph record (with the 1961 and 1965 versions) illustrating the 100 acoustigraphic displays of the article. This report and record were re-published as a portion of a book entitled *Newborn Infant Cry* (Truby, Bosma, and Lind 1965) which further contains an additional 100 cry-sound spectrograms and correlated X-ray cineframe tracings, and demonstrates the application of sound-spectrographic analysis to pediatric diagnosis at birth of the various pathological manifestations of brain damage. For, just as 'a voice' is hypothetically unique with reference to SPEECH (and regardless of the ostensible LINGUISTIC nature of the signal) so is any human voice also and at the same time indicative of a particular, complex, physiological — and/or pathological — *status quo*, whether the subject or patient is a neonate or at any other infantile PRESPEECH state as conventionally defined, or a 'beginning talker', or a linguistically developing child, or a linguistically mature adult, or at some stage of geriatric decline.⁷

Throughout the past fifteen years I have faithfully tape-recorded from birth — and DURING birth, by larynx auscultation — the prelinguistic and linguistic development of our own six children, three girls and three boys. The voice patterning and *idiomarks of these children, as seen on speech-sound spectrograms, cry-sound spectrograms, and the like, demonstrate longitudinal individualities so obvious that any member of this paper's audience could easily and consistently discriminate the children directly from the sound spectrograms.

For me this discrimination series has been an encouraging substantiation of the contention of my disclosure paper at the 1960 Meeting in San Francisco of the Acoustical Society of America (Truby 1960), and of my 1960-61 NINDB tour that the voice individuality noted generally in human speech is noticeable from — and even during — birth. At that Meeting I introduced the notion of infant-cry-sound indi-

⁷ Although neither Voice Recognition nor any aspect of individuation were pursued in the two studies noted in this footnote, it is a matter of record (and of phonograph record) that in the United States, more or less contemporaneously with our own American-Swedish-supported research in Stockholm (and quite independently and innocently of our work, as we of theirs), Drs Samuel Karelitz, pediatrician extraordinary, and Vincent Fischelli and group were also pursuing the significance to pediatric diagnosis of infant cry — in terms of other parameters and procedures than ours — under sponsorship of the U.S. National Institute of Mental Health (Karelitz, Karelitz, and Rosenfeld 1959), Karelitz and Fischelli (1962), and that in 1951, Arthur W. Lynip reported the first application of the sound spectrograph to tape-recorded 'birth cries', subsequent cries, laughs, speech-like sounds, speech-imitative sounds, and finally "speech sounds proper", sampling the initial 56 weeks in the life of a single, normal child (Lynip 1951). The only Voice Recognition research focused on physiological and pathological idiosyncracies of a NON-linguistic nature appears to be my own, the diagnostic aspect of which is the primary concern for the infancy studies. In this last regard, Pediatrician Arthur H. Parmelee, Jr., M.D., contributed a strong statement in 1962 of the significance to neurologic diagnosis of sophisticated studies of infant crying (Parmelee 1962).

viduality as well as the terms CRYPRINT and CRYPRINTING. In the same year, the above-cited article, "Cry Sounds of the Newborn Infant" (Truby, Bosma, Lind, and Karlberg 1960, Truby, Lind, and Bosma 1961, Truby, Bosma, and Lind 1965) introduced this notion variously and these terms in publication.⁸ (Except for a single, inconclusive but predictive, internal report of 1944 (Gray and Kopp 1944), the terms VOICEPRINT(S) and VOICEPRINTING and the notions conventionally associated with them didn't appear until 1962 (Kersta 1962) in conjunction with the "contour spectrogram" of Bell Laboratories and of Lawrence Kersta especially, who was interested enough to process, immediately in 1962, a tape recording of my 1960 phonograph record "Cry Sounds of the Newborn Infant" (Truby, Lind, and Bosma 1961), which contained both neonatal renditions and my own imitations of the universal types of newborn-infant cry sound.)

Now, when one begins to speak of the individuality of a given infant's 'baby cries', immediate notice will be taken by mothers, grandmothers, wardnurses, aunts, and others, and it was the ability of certain representatives of these categories to identify their own or specific infants that lent encouragement to Lind, Bosma, and myself as our study got under way in 1957 in Stockholm. But, as in speech and language affairs, only SOME of the mothers *et cetera*, have the facility to make consistent identifications. This I have tested right in the maternity wards, both from tape-recordings and with in-the-flesh babies, and I really didn't need any so-called 'controlled procedure' to demonstrate the obvious — some mothers could, and some mothers could not identify their own babies' cry-voices... out of 300 cases. The following statement about some of my marine animal subjects comes from a recently published article entitled, "Dolphins and Language" (Truby 1970b:182), and could be applied with equal impunity to human mothers and so forth:

Of course, as with *Homo sapiens*, we find *clever* dolphins and killer whales and porpoises and *not-so-clever* dolphins and killer whales and porpoises.

But no matter how insensitive some are to certain vocal performance, every human being — with only extreme pathological exceptions — is able to recognize certain 'voices', in the general sense. Little wonder, then, that individuals are called upon to testify in courts-of-law that this or that 'suspect' may be recognized or identified

⁸ The cryprints, term and concept dating from 1960 (Truby 1960, Truby, Bosma, Lind, and Karlberg 1960, Truby, Bosma, and Lind 1965, Truby, Lind, and Bosma 1961 — that is, the visual-acoustic characteristics of the crying — of a given neonate are, under a sufficiently detailed analysis, his and his alone; yet, being human produced, the consequent asymmetries make iteration impossible, or at least unlikely, and the observable result is that each cry sound is physically different from every other (p. 45) ... The sounds of cry are (normally) present from birth, easily accessible for permanent recording and analysis, and, in the cryprint form herein demonstrated, appear, so far, to be potentially as reliable and consistent as fingerprints where individual identity is the issue (p. 45 of the 1961 version, p. 46 of the 1965 re-issue). This last analogy was reiterated in January of 1961 in invited seminars at Bell Laboratories and Haskins Laboratories and in 36 other scheduled lectures during my 1960-61 NINDB tour, and many times in the intervening years. I have of course considered the differing physical natures of the respective identificatory phenomena. And see footnote 22 of this article.

on the basis of 'voice' characteristics. The only valid observation in this regard is that there are SOME people who should NEVER be so called-upon as identifiers, and that even the most 'talented' listener COULD make a misidentification, not only due to the grossness of LINGUISTIC demonstration generally, but to the subliminality of those individualistic features which are involuntary and immutable.

Something of the above is what I intended in my abstract when I wrote that one man's opinion is clearly NOT as good as another's in pattern recognition affairs. As a matter of fact, when it comes to *vocal*-pattern differentiation, one cannot help thinking of Henry Sweet, who could reputedly isolate 120 vowel continua, or Daniel Jones, or A. Lloyd James, or John S. Kenyon. On the same wavelength, Henry Lee Smith has on many occasions correctly labeled individuals from highly specifically pin-pointed speech localities, as have Miles Hanley, Alan Hubbell, Marshall Berger, and many others of our acquaintance. C.K. Thomas conducted over 10,000 personal speech-interviews while attempting to catalog so-called 'American English', and dialectology reverberates to the talents of such as Hans Kurath, Bernard Bloch, Raven McDavid, Fred Cassidy, and various other linguistic field workers, whether wearing the cloak of anthropology, ethnography, or linguistics proper.

As stated in the abstract, man has clearly been universally aware of the individuality of individual voice performance since at least the dawn of written history. The literature of every culture and of every age is punctuated with instances of this awareness... tales, plays, poetry, ballads, sagas play on it... as do the *Avesta*, the *Koran*, the *Old and New Testaments*,⁹ the *Veda*, and all other such accounts of man's imagined heritage. From the *Mahabharata* to Aesop's *Fables*, from the *Middle English Bestiary* and Sir Thomas Browne's *Pseudodoxica Epidemica* to the *Arabian Nights* and *Mother Goose*, from Boccaccio, Chaucer, Shakespeare, and on and on... one encounters man's consciousness of the role of 'voice individuality' in matters of identity confirmation and duplicity, of acceptance and rejection, of affection and hostility, and of intermediary and related circumstances. This 'voice individuality' identifies with livelihood... as to entertainment and salesmanship and personal promotion generally, and with life itself — as manifested in war, in civil security, and in the legal aspects of human conduct and affairs.

Next in my abstract appears the observation that INTRASPECIES and INTERSPECIES recognition — on the sound plane, of course — has also been demonstrated generally

⁹ In the hand-me-down language of *John* 10.4,5, we read: "A sheep does not mistake any other human voice for the voice of its shepherd, and a ewe in a large flock knows the call of its lamb from the call of any other lamb". These references have to do with VOICE — not so the oft-cited example also from biblical times *re* the PRONUNCIATION of representatives of different regions of the term reported as *shibboleth*, which is an instance of 'dialect identification' independent of 'VOICE recognition'. While it may be so that even the most so-called 'primitive' people are instantly aware of any deviations in the pronunciation or usage expected, the size and consistency of the community is a significant individuation factor in this dialectological consideration. This report hopes to call attention to the hierarchical levels operating in 'speaker recognition'...language, dialect, regional and/or social variant, intimate-community variety, idiolect, and finally: sublinguistic idiosyncrasy of speech-sound production and/or schedule.

throughout history and instrumentally since the advent of phonographic recording. As with accounts of man's awareness of 'voice individuality', the literature of most cultures treats generously the species-specific phenomenon of sound rapport throughout the animal kingdom, from the mightiest of mammals to relatively insignificant insects, with accountable exceptions all up and down the phyla. And corollary to this is the animal and human awareness of EXTRA-species signal. Simply put, living creatures generally and almost universally interrelate in the sound domain, utilizing species-specific signals in the recognition of sympathetic as well as UNSympathetic fellowcreatures. Within a particular species there appears to be signal transfer for such anthropomorphically analogous operations as Warning, Complaint, Exhilaration, Despair, and the like, reminiscent — again in human terms — of Communication *per se*. The more sophisticated students of animal behavior are presently quite concerned with definitions of 'communication' and 'language' and the like, in these challenging regards. Dolphins, for example, use entirely different MODES of acoustic signal for humans as contrasted with what they use for other dolphins. It was while comparing these different modes that the delphinologist John Lilly, (Lilly, Miller, and 1967, Lilly, Truby, Miller, and Grissman 1967, Truby 1967b, 1970b, Truby and Truby Lilly 1967) and I were able to posit entire networks of possibilities and plausibilities in the communication domain. These networks exposed at once the excruciatingly complex nature of human communication and linguistics and the inadequacies and inappropriatenesses of conventional language and speech analysis for considerations of the highly complex Bottlenose Dolphin, for example.

And ACROSS species, it is commonly observable that signals from an aggressor are generally rapidly processed by most defending species in time to bring about flight and/or other defense tactics as relevant. The relatively minimal — when compared with human — acknowledgment of the human 'voice' in the nonhuman-animal world has, of course, also been documented historically, but it is scientifically dangerous indeed to bandy about such notions as Language, Speech, Voice, Communication, and even Mimicry... when treating non-human-animal sound-performance, whether quasi-vocal itself or as an apparent response to human-generated signal. In all events, phenomena observed by humans in the pre-sound-recording days have been corroborated or refuted or tempered, as the case might be, since the advent of phonographic recording. At least now, in the words of my aforementioned 1960 article "Cry Sounds of the Newborn Infant" (Truby, Bosma, and Lind 1965, Truby, Bosma, Lind, and Karlberg 1960, Truby, Lind, and Bosma 1961), p. 14:

Sound recording, and especially magnetic tape recording, gives the investigator or clinician permanent storage and thus the opportunity to listen repeatedly to the *same sound(s)* — not merely *similar* sound(s) treated by the listener *as* identical, but precisely the *same* acoustic performance(s). [It has been known for 25 years that] it may be readily demonstrated with a sound spectrograph that no sound — and of *course* not even 'a *speech* sound' — can be willingly or even accidentally produced twice...even by the same speaker. Infant cry sounds [may at first seem] less variant, but clearly, no two infants cry alike, and no infant cries the

same way twice. Our permanent acoustic records of the past 14 years verify these observations (within the limits of our data corpora).¹⁰

Now, every reputable phonetician spanning the years from Henry Sweet and Edward Scripture to the present moment has posited the speechsound uniqueness just noted, but it was sound-spectrographic analysis and advances in oscillography which confirmed these postulations physically. For example, Scripture, in 1902, in his 653-page(!) book *Elements of Experimental Phonetics* wrote (1902:118):

A speech sound produced by an individual is the result of a very large number of fine adjustments of the speaking apparatus influenced by an infinitude of past and present experiences in hearing, thinking, and speaking. The sound varies from moment to moment and from one occasion to another. With sufficiently accurate methods of measurement no two sounds would be found alike;

Scripture's forward-looking assertion materialized, some 40 years later, in the form of sound-spectrographic analysis, as discussed, which operation is presently adequate, IN THE PROPER HANDS, to the specification of idiosyncratic particulars demonstrably present in the speech signal. The startling fact is that 70 years ago Edward Wheeler Scripture touched upon almost every fundamental concern of the experimental phonetics of his day as of our own day... including numerous observations about perception and behavior presently being painfully rediscovered by psychology. It seems clear that were the "speech engineer" of today truly conversant with the contents AND THE IMPLICATIONS of this very thorough book of 1902, that engineer or phonetician or what-have-you could more often than not throw away his divining rod and spend his time actually digging! Scripture wrote, as Jules Verne wrote on other sources, of a time to come... and of a time that DID come.¹¹

To me, Voice Recognition implies the ultimate in phonetic analysis — *la phonétique par excellence*, as it were, and reference to phonemics is of secondary — if not minimal

¹⁰ (p. 11 of the same article: an issue of this report is, as stated, a classification of the different kinds of vocalization common to cry. ...to devise an OBJECTIVE method for the acoustic analysis of infant cry, since there does seem to be a significant relationship between cry and neurophysiological state. (p. 10) ... Auscultation by microphone and/or by stethoscope, DIRECTLY AT THE LARYNX, is highly rewarding and discloses acoustic activity quite imperceptible to 'the naked ear'...It will be demonstrated, primarily by deduction from the visual patterns seen on sound spectrograms and on oscillograms, that the PERFORMANCE pattern of cry is exceptionally individual — very possibly the infant's most distinctive motor activity...(p. 11). Conclusions: As a human-produced phenomenon, the cry of the newborn infant is uniquely individual. (p. 45)

¹¹ It has been stimulating to me to hear during this Congress that I am neither alone in my concern over the widespread ignoring of predecessorial research and publications, nor in my esteem for Scripture and others as noted. In the former regard, any contributor to such a Congress as this owes it to the general professional stature to be thoroughly familiar with at least the Proceedings of all previous Congresses and to present material in the light of that information, witness the misgivings expressed in the Opening Remarks by Bertil Malmberg, as in the critique of others of our number present. In the latter regard, as was done for Helmholtz (1877, 1954), the entire 1902 edition of *Elements of Experimental Phonetics* is being reissued under the aegis of and with a new introduction by our attendant colleague John W. Black, and is presently in press. Scripture continued his cogent contributions to the acoustic phonetics literature for over 30 years, e.g., Scripture 1933.

— significance to Voice Recognition proper, and is mentioned here chiefly for the sake of indicating the deceptiveness inherent with the PHONEME CONCEPT, and the increasingly general misinterpretation of the phoneme as 'a sound' and its consequent inappropriateness frequently for constructive application to Voice Recognition procedures and concerns. Idiolectalism is more to the point than Phonemics broadly, though here again, it is both the specific and the consummate PHONETIC — to the ignoring of PHONEMIC — considerations that have relevancy for Voice Recognition. This is, in fact, the beginning of the indication of the dichotomy cleaving SPEECH and VOICE. By common definition, Voice Recognition is concerned, for human affairs, with ascertaining the speaker's identity by whatever relevant operation,¹² but considerable investigatory energy has been expended fruitlessly — and more often than not, misleadingly — in unnecessary attempts and efforts to find so-called PHONETICALLY identical (or otherwise LINGUISTICALLY identical) contexts. As anyone with much experience in looking at speech-sound spectrograms should be able to attest, and as has been brought out in the literature and as recently as the paper presented by Professor Stevens at this Congress (to be found on pp. 206-232 of this volume) and some of his discussants, the *same individual* may, in many instances, demonstrate greater variety of performance in the vocal productions of a given utterance than may different individuals from the same linguistic community, whether that utterance is found in *different* contexts or in apparently *identical* contexts. But, a given individual will — at least hypothetically — stamp any and all of his or her vocal output — 'linguistic sames' or not — with idiosyncratic PHYSICAL FEATURES which, though not apparent to even the most highly 'trained' ear, will nonetheless be apparent visibly with optimal speech-sound spectrography! And once these *idiomarks are manifested for an investigator, they will obviate any requirement of 'linguistic sameness' for Voice Recognition purposes. In short, neither homophones nor homonyms by whatever definition are either necessary or consistently useful in attempts at speaker identification.¹³

¹² Edward Sapir, in 1927, had indicated the significance of speaker identification in the differentiation of individual and society: "society has its patterns, its set ways of doing things...while the individual has his method of handling those particular patterns of society, giving them just enough of a twist to make them 'his' and no one else's." (1927:892-4). Paul Garvin and Peter Ladefoged (1963) "would like to elaborate this conception of Sapir's" and contend that "human beings are indeed capable of recognizing not only individuals but classes of individuals by voices", that there is "clear differentiation of the two basic types of information which are contained in the voice signal: speaker identification and message identification", that "speaker-diagnostic characteristics...are fundamentally due to organic factors...susceptible to modification by learning...idiosyncratic speech patterns" reflecting "social and/or regional varieties of speech by a given individual", and thus that "speaker identification" is a *speech-habit* consideration rather than a *voice-idiosyncrasy* phenomenon. Speech — or idiolect — is independent of VOICE. See also footnote 17 herewith.

¹³ Strictly speaking, the term HOMOPHONES means 'same sounds', HOMOGRAPHS means 'same spellings', and HOMONYMS means 'same names' — i.e., 'having the same sounds *and* the same spellings' with no morphosemantic indications of other restrictions or specifications. However, the most useful classification restricts *homophones* to those words or other linguistic elements which 'sound alike' but differ in spelling and meaning, *homographs* to words or other linguistic elements which 'look alike' — i.e., are spelled alike — but differ as to sound and meaning, and *homonyms* to words

Voice Recognition by animals, in animals, and of animals, is a matter more closely related to *human* intraspecies voice recognition than is realized, and I am impelled to report that my present attitudes about language and communication owe a lot to my intimate work during the past six years with the Bottlenose Dolphin, *Tursiops truncatus*, and with the very badly mis-named 'killer' whale, *Orcinus orca*.¹⁴ These nonhuman associations have opened my eyes and my ears to a phenomenal sector of the world of sound in which we live much too innocently. In trying to work out a valid *modus operandi* for the examination and study of communication, both for extrahuman intraspecies and for human-anchored interspecies concerns, I soon saw how unsatisfactory is much of our phonetic and other applied-linguistic dogma, doctrine, hypothesis, and lexical and semantic inventory appraisal... not only for interspecies considerations of whatever nature, but even for unambiguous human-linguistic considerations and procedures. In a paper delivered at the X International Congress of Linguistics in 1967 in Bucarest, entitled "Language and Dolphins" (Truby 1967b), I made a few statements particularly relevant for certain concerns of the present paper:

In an alien species, often only the *sound* is accessible to investigation. We must thus look to acoustic considerations alone for elucidations of communication...In the case of *Tursiops truncatus*, the Bottlenose Dolphin, the generator is non-human, and neither subject to human limitations nor bound by human conventions. In such an instance, it is perhaps better to proceed without reference to linguistic criteria, especially where based on conventional human language. E.g., *Tursiops truncatus* utilizes signals in his interspecies application which differ extremely from the speech signals of *Homo sapiens* as to manner of production, general and specific anatomy involved, motivation, and at the same time, as to acoustic composition and frequency range [operating as high in spectral frequency as 200 and even 300 kHz]. And physically differing basic signal types are utilized for respectively differing general situations and circumstances.

Voice Recognition by MACHINE implies not only the application to analysis procedures of the sound spectrograph, oscillograph, and other *acoustigraphic instrumentation, but also the development of computer tactics adequate to the horrendously complex task of sound-pattern discrimination. This accomplishment will be the result of an intelligent marriage between Communication and Communications experts.¹⁵

or other linguistic elements which 'sound alike' and 'look alike' but differ in meaning. This distinction of *homophones* and *homonyms* has relevance for Voice Recognition, since the manifestation of **homonemic* differentiation sometimes reflects subphonemic idiosyncrasy, for which see Truby 1964b.

¹⁴ I say 'misnamed', since he is no more or less a 'killer' than any other predator and has evidenced a gentleness in his interaction with man, woman, and child that is unbelievable in an animal ranging from 20 to 30 feet in length and from 4,000 to 8,000 pounds in weight, as seen in slides I made of Hugo, local *Orcinus orca*, and my school- and preschool-age children and me, all in obvious congenial rapport with said — or missaid — 'killer whale', actually the largest of the 55 species of *Dolphin* which constitute the Family *Delphinidae* (Montagu 1962) as will be discussed further, along with a plea for finding out what a 'porpoise' is before calling dolphins 'porpoises' (Truby 1970b). And see footnote 23 herewith.

¹⁵ 'Communications', by the way, is not at all, it so happens, the plural of 'Communication', nor

In the previously mentioned paper presented in 1960 at the 60th meeting of the Acoustical Society of America in San Francisco (Truby 1960 and 1967e:541), I indicated some of the individualistic visible-acoustic parameters functioning in human voice manifestation on sound spectrograms, and I listed these visible correlates in terms of seven, general, directly measurable DESCRIPTIVE FEATURES. These descriptive features had been reported earlier in a compendious book (Truby 1959) in which I published 1100 sound spectrograms illustrating, in visible-acoustic terms, many gross and specific physical differences (presently being 'rediscovered'), such as duration and nasalization among 'phonemically same' vowels, devoicing among presurd clustered nasals, the physical simultaneity of initial-cluster coarticulations, the incompatibility for phonological evaluations of 'phonemically same' — but phonetically distinctive — sound continua, and so on and so forth. I made these and other observations during the first four years of Fant's Speech Transmission Laboratory (Taltransmissionlaboratoriet), which followed the four years at Haskins Laboratories, which followed the previously mentioned 3,000 spectrograms of my own speech sound. And as further preparation for looking at sound spectrograms, during my eight years of residence and research in Stockholm, I made an extensive independent, sonocineradiographic investigation of speech production, not only in 22 American English speakers, but in Spanish, French, German, Arabic, Finnish, and Swedish speakers as well. I think I learned more about speech-sound spectrograms in correlating them with the X-ray sequences than from any other laboratory procedure. From these observations my set of descriptive features evolved. As I reported at the time, the analysis of any speech-sound continuum in terms of these descriptive features will generate a relevant, continuous series of physically contiguous micro-

patterns which, since 1958, I have termed PHONETTES.¹⁶ These micropatterns should not be confused with the macropatterns once ambiguously approximated with the notion of "spectrophones" and "spectrophonemes" (Truby 1962b).

Phonettes, by contrast, are physically measurable sound-sectors and provide a workable basis for the scientific application of the sound spectrogram to voice recognition by man and, by extension, by machine. They should not in any way be confused with or concerned with phonemes — however you might define phoneme. [In this last regard, see Truby 1959, 1963, 1964a, 1964b, 1965d, 1967d, 1967e, 1967i, 1968, 1969, 1971a, 1971b, 1971c.]

Kersta writes of "96-millisecond time windows which are incremented at 16-millisecond intervals over the entire word spoken" (1969), and I feel certain that this and similar time-sampling procedures reported over the past ten years or so could accomplish what I intended with phonettes, which I described in 1959 as descriptively minimal segments

referable to (but not coincident — in any sense — with) [either phone or] phoneme... definable expressly on the basis of pattern-change...determined entirely by eye...phonic, though their classification is visual... It was, and remains, conceivable to me that *phonettes* could be determined *mechanically* (which *phonemes* cannot!), without recourse to linguistic criteria.

Such specification of physical particulars is needed, not only for indicating dialectological particulars but also for quantifying individual VOICE characteristics. I myself, for instance, have been indicating the nature of these idiosyncracies to my acoustic-phonetics classes for the past seventeen years, at the Universities of Kiel, Uppsala, Stockholm, California at Berkeley, and Miami, and variously at Haskins Laboratories, the Speech Transmission Laboratory in Stockholm, the IBM San Jose Research Laboratory, the National Institutes of Health, and Communication Research Institute, with each of which laboratories I was associated for periods of three years or more. Nor would I ever have embarked on the long-range infant-cry-sound studies reported above had I not been conversant with these NON-LINGUISTIC particulars of sound-spectrographic analysis,¹⁷ and I would certainly never have left a warm

are the two terms technically interchangeable, the most general usage error being the substitution of the former for the latter. The two disciplines themselves rarely ever overlap or are even very often compatible. As I stated in the Martinet *Festschrift* in an article entitled: "A Definition of Speech-sound Analysis, 'Speech Synthesis', and Speech", the terminological distinction communication/communications sets apart "the science of message-transmission from the *technological instrumentation* of such transmission, the former, communication, being the province of linguistics, neurology, zoölogy, biophysics — the latter, communications, being strictly an engineering concern" (Truby 1967e). For a succinct and cryptically frank critique of one of the least responsibly treated aspects of Voice Recognition, I recommend John Pierce's 1969 derogatory letter: "Whither Speech Recognition?" (Pierce 1969), with such pointed laconicisms as: "Speech recognition has glamor. *** Results have been less glamorous. *** People who work in the field are full of innocent (in their own view) enthusiasm. *** We all believe that a science of speech is possible, despite the scarcity in the field of people who behave like scientists and of results that look like science. Most recognizers behave, not like scientists, but [sic!] like mad inventors or untrustworthy engineers. The typical recognizer gets it into his head that *he* can solve 'the problem'. *** We would expect a scientist to check the literature concerning ideas, schemes, or information [and] devise a clear, simple, definitive experiment. *** The typical recognizer will have none of this. He builds or programs an elaborate system that either does very little or flops in an obscure way. A lot of money and time are spent. No simple, clear, sure knowledge is gained. The work has been an experience, not an experiment". This is followed by 40 references which "trace in detail the uncertain course... of speech recognition through the literature". In this juxtaposition these titles read like a spoof of science or a hocus pocus.

¹⁶ Descriptive Features (Truby 1959)

- | | |
|--|--------------------------------|
| 1. duration | (in terms of) |
| 2. absolute overall-intensity | 1. msec |
| 3. frequency limits and/or components | 2. relative dB |
| 4. source | 3. cps (Hertz) |
| 5. motion or lack-of-motion | 4. periodicity &/or turbulence |
| 6. rate-of-change of formant frequencies | 5. slope of F _n |
| 7. type of onset & termination (if applicable) | 6. cps (Hz) per msec |
| | 7. abruptness or smoothness |

¹⁷ With newborn-infant cry-sound, of course, one does not encounter linguistic criteria, and, in fact, I shake a reproving finger at those who have made "phonetic transcriptions" of infant cry-sound as I said in my Prague paper (Truby 1967d) in no uncertain terms. As one observes infants crying, one can imagine all sorts of linguistic performance, but, as I commented the other morning at Adrian Fourcin's paper (delivered at this Congress and to be found on pp. 48-62 of this volume), the infant mouth during cry is typically in the classical 'square-lipped' position, the tongue is in a relatively fixed involvement gesture with the tip and blade-edges cupped up, the entire tongue is in repetitive muscular agitation, and any articulation is accomplished with the mesopharyngeal

bed at IBM to track dolphins and other whales in icy ocean waters, had I not accumulated a large measure of confidence in the proposed analysis instrumentation. That is, I felt that while it was comfortable to continue to embellish the artifacts of instrumental 'analysis by synthesis', the investigation of another species was bound to throw light on linguistic evaluation as a whole.

As a consequence of broadening the objective and the scope of my own recorded-voice archives, I now have over 200 paranatal or birth-cry sequences, over 300 neonatal sequences, fifteen years of follow-ups of some of the birth-cry babies; over 100 different foreign languages utilizing the 'semantically same' three-page text; the essentially endless recordings of a dozen or so different Bottlenose Dolphins, and likewise for two marvelously vocal and cooperative 'killer whales'; an assortment of isolated, natural-environment records of bear, elephant, humpback whale, sea lion, harbor seal, carabao or water buffalo, and others; and assorted dogs, pups, cats, kittens, ponies, frogs, mice, and several species of both birds and bees!

I regret that time will not permit me to supplement the preceding excellent discussion of Ken Stevens (paper delivered at this Congress and to be found on pp. 206-232 of this volume) with my own list of the snares and traps and illusions both visual and auditory that speech-sound analysis is fraught with. In any case, I will venture a definition of 'the individual voice' as being A HYPOTHETICALLY UNIQUE COMPLEX OF PHYSICALLY MEASURABLE VOCALIZATION FEATURES. And while it may be, at least at the moment, that these reflections of individualistic vocal gestures are not necessarily readily apparent on either the auditory or visible-acoustic planes, I can assure you they are present nonetheless. So far I have not encountered duplications in adult speech-sound, children's speech sound, infant cry-sound, or dolphin 'vocalization', in these areas of my most multiplicitous data.

When I accepted the invitation of Professor Rigault to address myself to the issue of this Plenary Session, I did so with the feeling that I was invited here partially — though I hope not exclusively — to defend Lawrence Kersta's position in the 'voice-printing' issue. I was certainly — at one tense moment — single-handedly responsible for squashing an organized attempt to publicly censure Kersta professionally, and on another occasion publicly outspoken at defending Kersta's right to be treated as 'innocent' until proven 'guilty', and I do myself believe that the positive identification of speakers from optimal sound spectrograms is possible, with certain provisos and restrictions which I've considered obvious from the first. I am still startled to find a general innocence of the distinction between WHAT was said and HOW that particular what was produced! The very reason that it remains difficult to linguistically decipher

dorsal wall, which is labile in the newborn, and NOT with the organs or articulations implied by the traditional transcription, based, as it is, on idealistic ADULT speech performance. In fact, any reference to human language is demonstrably inappropriate, and direct observation of the infant during sound production reveals that a newborn-infant cry-sound HEARD as 'beginning with /w-/ or /m-/' did not involve 'labiality' in its production at all, the entire articulation being accomplished essentially pharyngally! The representation of these articulations in traditional 'phonetic transcription' on the basis of resemblances to adult articulations is irrelevant and artifactitious.

a 'mystery' sound spectrogram is that there are no comprehensive studies of How the speech of individuals manifests itself on sound spectrograms.¹⁸

The gist of the "Voiceprint" matter is that Kersta, after 35 years' experience with sound-spectrographic analysis, launched an independent venture to apply sound-spectrographic analysis as supporting evidence in legal proceedings. As a consequence he was called into courtrooms as an 'expert witness'. To some of his professional colleagues (Ladefoged and Vanderslice 1967:12-13) all of this was impalatable, and a pointed action was taken to force Kersta from such activity. My own objection to this censuring was — and is — that not only is it highly improbable that one scientist is able to ascertain or assess anyone's perceptual abilities, but certain of the things Kersta's critics said, for instance in an informative — but unintentionally misleading — 1969 *Science* article (Bolt *et al.* 1969), could with greater justification be applied to the present, traditional, and universal acceptance in courts-of-law of 'speaker identification' (by ear alone) by practically anyone who 'comes down the pike'!¹⁹ I believe that my own provisos in voice identification for such investigations include unimpeachable safeguards for validity, and some of these have also been put forth in the essentially fair and cautious *Science* article (Bolt *et al.* 1969). Though I have so far successfully avoided the witness chair, I too, in recent years, have been engaged professionally in 'voiceprinting' regards — by prosecuting attorneys, defense lawyers, radio and television companies, and private individuals, but I have an open set of definite guidelines according to which I make judgments in these issues. Had I needed reassuring, the recent Michigan police study by Tosi *et al.* (1971), is certainly corroborative of my long asserted optimisms about voice individuality.²⁰ And

¹⁸ And when Ken Stevens says more attention is being paid to glottal source and that sort of thing — in *individuals*, that's on the right tack. When he stresses *more measurements*, that's on the right tack. Phonettes: 1958 — that's not only on the right tack, but hitting the tack right on the head! (Truby 1959).

¹⁹ To accept the testimony of any ordinary 'witness' that he or she, three or four or six months ago or longer, remembers a particular voice as having said, 'Stick 'em up!' at The First National Bank or at the corner of Third and Maple...is ridiculous! Yet, it's still being done — and accepted in courts of law, at least as testimony to be *considered* by jurors or judge. Certainly, an approach through some kind of instrumental analysis, under professional conditions, is — in the right hands — a wiser and more *valid* approach.

It is certainly highly regrettable that the 'hysteria' of Ladefoged and Vanderslice (1967) and related 'reports' should ever have been permitted to reach such 'Chicken Little — The Sky is Falling' proportions as is implied from the list of 'defenders' named on pp. 12-13 of their article. Such conduct is a worse social ill than the alleged 'scientific misconduct'. A strong but friendly word of caution was all that was called for. I am in complete concurrence with Ladefoged's implication that dialectological or idiolectological criteria are inadequate to the proposition of speaker individuation — they must be validated with sub-linguistic criteria as discussed here in connection with footnote 22 and elsewhere. *I herewith so caution everyone concerned with "speaker individuation"!*

²⁰ "34,996 identifications of 250 speakers were attempted by 29 trained examiners, *solely on the basis of the inspection of sound spectrograms*, in tasks involving contemporary and non-contemporary utterances in closed and open trials of 9 or 6 clue words uttered in isolation, in a fixed context, and in random contexts, etc. Results conclusively confirmed Kersta's original experimental data and extended the range of trial instances in speaker identification reliability, also accumulating further information on both..." (Tosi *et al.* 1971:34).

others are currently voicing similar optimism.

But, they're still dealing with SPEECH-SOUND and how SPEECH is manifested... according to LINGUISTIC strictures. They're talking about *dialectological* particulars — not, about the voice *per se*.²¹

In an internally published article (Truby 1970a): “The Validity of ‘Voiceprinting’, ‘Cryprinting’, and other ‘Soundprinting’”, I responded to the entire issue. Certainly, *some* good has come out of the matter as a whole, since long-overdue scrutiny has begun to be given sound spectrograms. But, as I wrote in my abstract, there is nothing more subjectively personal than the interpretation by *one* individual of *another* individual's performance or knowledge or ability. As in most human affairs, one man's opinion is not necessarily as good as another's.

[At this point, a series of black-and-white and color slides was presented, in individual and collective support of my position that *voice* individuality can be demonstrated with sound spectrograms, beginning with the earliest neonatal cry-sounds of individual infants (actually paranates, as described below), continuing into and through PRE-speech vocalization, into the beginnings of and throughout the DEVELOPMENT of speech, and into the entire range of matured speech. For this, slides of representative sound spectrograms, some correlated cineradiograms, and other acoustigraphic displays were presented.

[Specifically, idiosyncracies of acoustic output *for the same individual* were demonstrated for... paranates (larynx auscultation of the emerging infant during birth — vocal efforts leading to ‘first cry’)... neonates (universal cry-pattern types in terms of individual performances)... infants (during pre-linguistic vocalization)... children (during unmaturation speech)... and adolescents (during matured speech), thus indicating the high degree of individuality in terms of which individuals may be charted and differentiated from birth through childhood and so on.

[An example of the degree of specificity possible during visible-acoustic examination (‘in the right hands’) was offered in terms of a slide of a sound spectrogram of an utterance of ‘glanced’ (mixed transcription: [glæŋtstʰ]) demonstrating idiosyncracies involving not only such general phonetic features as nasalization of the vowel in anticipation of “a nasal continuant”, itself DEVOICED, however, in anticipation of ‘a voiceless sector’, mandatory final consonant explosion, etc., and even the individualistic — and thus, idiolectal — *degree* of vowel nasalization in terms of the articulation-sequence timing pattern, the degree of initial consonant pre-voicing and voicing, the relative duration of (in this case, a ‘silent’) nasal continuum preceding voicelessness, the particular ‘kind’ of lateral here clustered, and other observable — and thus,

²¹ It is undeniable that the speech — or idiolectal — aspect is of a certain individual significance, but the individual *voice*, though a frequent parameter of idiosyncrasy, is not necessarily germane to *idiolect*, and especially vice versa! Put hierarchically, ‘voice recognition’ is very likely impossible if the ‘recognizer’ is unfamiliar with the speaker's *language*, more probable with familiarity with the speaker's *dialect*, even more probable with familiarity with the speaker's *idiolect*, and the most plausible where there is familiarity with all of these plus the speaker's *voice* idiosyncracies.

measurable — acoustic manifestations of a definitely subphonemic significance, but also instances of pattern particulars *not* attributable to dialectal, idiolectal, homorganic, assimilative, or other phonetic relationships, but to the particular set of supraglottal conformations of the individual, as a response to his particular neuromuscular (i.e., feedback-directed articulatory and source-generated) scheduling.²²

[Other slides supported remarks about research with dolphins and whales, the taxonomy comprehending the 22 genera and 55 species of sea mammal in the Family *Delphinidae* (Montagu 1962, Truby 1970b), the popular confusion about ‘dolphins’ and ‘porpoises’²³ the attempts at “communicating” with dolphins, the training of a so-called ‘Killer Whale’ (see footnote 14 herewith) *Orcinus orca* — the largest species of *Delphinidae*, and demonstrations of in-the-water rapport with such mammals.

[And some slides treated early and late ‘synthetic speech’ research and instrumentation, as related, in all instances, to Voice Recognition considerations.]

Author's note: I have tried, throughout this report, to indicate — even at the risk of being accused of immodesty — something of the complex of associations and investigations which has generated in me such a high measure of confidence as regards individual-voice uniqueness or voice individuation. The amalgamation of basic and applied research, in sound-spectrographic analysis and sonocineradiography of neonates, infants, children, adults, and dolphins and in synthetic speech and phono-

²² It is these features which specifically validate the comparisons with *fingerprinting* implicit in such terms as *cryprinting*, *voiceprinting*, and *soundprinting* in general — the *idiolectal* features corresponding to “the gross types of ridge patterns, such as loops and whorls... used for classification and indexing; these types are determined mainly by heredity and thus have only limited power in differentiating persons”, and the *idiosyncratic* features corresponding to “The fingerprint features... ultimately used for identification — the most minute details of the skin ridge patterns such as bifurcations, terminations, and interruptions”. (Bolt *et al.*, 1969, as modified from Cummins and Midlo 1961, Galton 1892). What holds for skin holds point-for-point for voice, as further almost directly quoted: “These details are determined mainly by random processes in prenatal development. There are a [sic!] sizeable number of these minute anatomical features [for each individual]. There are [sic!] an enormous number of possible combinations of these features...”. It is regrettable that this essentially cautious and considered article (Bolt *et al.*) is weakened by such misimplications as “changes with growth and environmental influences could be expected”, “Whereas fingerprint patterns cannot easily be faked or disguised, a speaker can learn to alter his voice and imitate, with some success, the speech of other persons”, and “In view of basic differences between fingerprints and voice patterns and the inherent complexity of spoken language, we doubt that the reliability of voice identification can ever match that of fingerprint identification”. These three last cited statements are either erroneous, non sequitur, or immature, in the light of present biophysical, bioacoustic, and broadly biolinguistic sophistication related to the questions of *soundprinting* and interpretation (cf., footnote 8 herewith), taken comprehensively. Thus, in spite of the value of their recommendations for generally establishing “know-how” in this concern, I consider the subtitle “How do scientists view its reliability for use as legal evidence?” of Bolt *et al.*, misrepresentative, misleading, and journalistic in its most conventionally plaguing sense. And see Truby 1970a.

²³ All 55 species being *dolphins*, some six of which beakless, small-brained, small-sized, easily differentiable species are rightfully *porpoises*; thus all true porpoises (the mammals) are dolphins, but 49 species of dolphins are NOT porpoises, and the 9-foot, 500-pound, bottlenosed species *Tursiops truncatus* found performing in seaquaria and on television has a brain weight of 1,700 grams, as contrasted with the 5-foot, 200-pound, beakless harbor porpoise *Phocaena phocaena* with a brain weight of but 400 grams.

logical alphabetistics (Truby 1958-1971, Truby *et al.*), has alerted me to the prospects recited in this paper and has made it impossible for me to accept the phoneme as "a sound" or as anything physical or such concepts or notions as *speech sounds*, diphthong, "blend", "simple vowel", "long vowel", "short vowel", and other unsupported — however traditional — terminological paraphernalia (Truby 1959, 1963, 1964a, 1964b, 1965d, 1967d, 1967e, 1967i, 1968, 1969, 1971a, 1971b, 1971c).

I do take this opportunity to thank Emerita Professor of English, and presently Adjunct Professor of Anthropology, Louise W. Hanley for editorial assistance and commentary, a portion of which latter is introduced here for its relevancy:

The author gives an account of his investigation of Voice Recognition, in which he has not depended upon a *single* line of study but has approached the work from many different angles and points of view. The results of his work, up to this point, are now available — an example is his book published in 1959, in which we find: "The resultant parametric sets, or *phonettes*, indicate a physical framework for the scientific application of the sound spectrogram to individualistic voice recognition by man — and, in extension, by machine... the relevant computer program being but a hypothetical step away." In other words, the work has been carried to the point where there is a definite procedure for Voice Recognition by the human auditor; from that point to recognition by computer is a comparatively short step.

CONCLUSIONS

As a consequence of my involvements and/or in summary, I do believe that no two voices are physically identical... or thus, ACOUSTICALLY identical... and that we either have or will soon have sound-analysis 'hardware' adequate to the task of demonstrating this individuality, from infancy through adulthood. And, in this regard, I believe that a newborn infant's identity can become a matter of physical record on the basis of the relevant cry-sound classification... that the CRYPRINTING of infants will come to have every bit as great a reliability as fingerprinting, palm-printing, and footprinting, for examples, in spite of physical considerations ostensibly to the contrary.

Directly related to this, but more importantly, I believe there is a diagnostically significant network of cry-pattern for a given infant... comprising the birth-crying proper, the transitional crying during respiration establishment especially, and the subsequent neopostnatal crying once the infant's respiration is independently and reliably established. From this network of cry-pattern the neurophysiological-neuropathological state of every infant is ascertainable. Put simply, the way an infant tries to start crying, starts crying, and cries once breathing independently should ultimately indicate an entire spectrum of disability, from severe brain damage on down. This may be extended to discussions of *linguistic* identity, as a further speculation.

And, still related to the notion of the acoustic individuality of voice, I believe that in spite of the complexity heaped onto the analysis by linguistic — in the sense of dialectologically phonemic — features, we will come to a refinement of acoustic

analysis which will move the reliability of VOICEPRINTING — as conventionally defined — ever closer to the reliability of CRYPRINTING, as described above. It must be clear that CRYPRINTING is based on each infant's *involuntary* exploitation of the particular vocal mechanisms and should directly reflect anatomical and neuromuscular individuality. VOICEPRINTING, by graded contrast, if restricted to speech manifestation in adults, implicates the imposition of linguistic pattern-conditioning onto the individualistic anatomical and neuromuscular performance. I believe the necessary analysis refinement will complicate — and simplify — our scientific future.

I further believe that nonhuman-animals demonstrating highly complex neurological circuitry are also specifiable as to species-specific 'vocalization' uniqueness, and that once we are scientifically adequate to the task of isolating the relevant criteria, we will open communication doors only isolatedly 'dreamt of' at present. If my optimism in this area is supportable, our knowledge will burgeon throughout the animate world, as we peel away the crusts and layers of ignorance presently hiding a prospectively veritable multitude of communication phenomena.

My final complex belief is that our communications engineers are capable of developing instrumentation and programming which will relieve the COMMUNICATION scientist of most of the burden of physical analysis based on the interpretation of acoustic features, once the body of communication data is evident.

Having placed these optimisms before you, I look forward to your criticism, and I thank you for your attention and your patience thus far.

Language and Linguistics Research Laboratory
University of Miami
Coral Gables, Florida

WORKS CITED

As noted in the second paragraph of this paper, and as implied with "Works Cited", the following is in no way exhaustive, either as to references or to personal bibliography. Much if not most of the literature treating "speaker identification" would be irrelevant here, since it is not VOICE oriented. A number of crysound studies are cited, since an aspect of "cryprinting" is indeed relevant to Voice Recognition considerations. The "pseudospeech" references, as the PHONEME references, are of relevance to the caricature nature of speech — and thus, language — perception. For further discussions of relevancy, see footnotes 2, 3, 9, 12, 17, 19, 21, 22 of this article, and elsewhere.

Bell, A.M.

1867 *Visible Speech: The Science of Universal Alphabets: Self-interpreting Physiological Letters, etc.* (London).

Black, J.W. [see footnote 11 herewith].

Bolt, R.H., F.S. Cooper, E.E. David, Jr., P.B. Denes, J.M. Pickett, and K.N. Stevens

1969 "Identification of a Speaker by Speech Spectrograms — How do Scientists View its Reliability for Use as Legal Evidence?", *Science* 166:338-343.

Bosma, J.F., H.M. Truby, and J. Lind

1962 *Cry Motions of the Newborn Infant*, 44 cry-sound spectrograms and correlated X-ray

- cineframe tracings and inverse plethysmograms (Stockholm). (Also in Truby, Bosma, and Lind 1965:61-92).
- 1963 *Studies of Neo-Natal Transition: Correlated Cineradiographic and Visual-Acoustic Observations*, 43 cry-sound spectrograms and correlated X-ray cineframe tracings of projected 35-mm negative (Stockholm). (Also in Truby, Bosma, and Lind 1965: 93-109).
- Cummins, H. and C. Midlo
1961 *Fingerprints, Palms, and Soles: An Introduction to Dermatoglyphics* (New York, Dover).
- Galton, F.
1892 *Finger Prints* (London, Macmillan). [facsimile reprint: New York, Da Capo Press, 1965].
- Dudley, H.
1936 "Synthesizing Speech", *Bell Laboratories Record* 15:98-102 (New York).
1939 "The Automatic Synthesis of Speech", *Bell System Monograph B-1169* (National Academy of Sciences, New York).
- Dudley, H.W. and O.O. Gruenz, Jr.
1946 "Visible Speech Translators with External Phosphors", in *Technical Aspects of Visible Speech* (= *Bell System Monograph B-1415*) and *Journal of the Acoustical Society of America* 17:62-73.
- Fourcin, A.J.
1971 "Perceptual Mechanisms at the First Level of Speech Processing", *Proceedings of the VIIth International Congress of Phonetic Sciences* (Montreal), to be found on pp. 48-62 of this volume.
- Garvin, P.L. and P. Ladefoged
1963 "Speaker Identification and Message Identification in Speech Recognition", *Phonetica* 9: 193-199.
- Gibbs, F.A. and E.L. Gibbs
1944 *Atlas of Encephalography* [current printing: Cambridge, Mass., Addison Press, vols. 1 and 2, 1950 and 1957].
- Gray, C. and G. Kopp
1944 "Voiceprint Identification", Unpublished Bell Laboratories report:1-14 (New York).
- Helmholtz, H.L.F.
1954 *On the Sensations of Tone, as a Physiological Basis for the Theory of Music*, 2d English ed. (New York, Dover). [translated from 4th German ed., 1877].
- Joos, M.
1948 *Acoustic Phonetics* (= *Language Monograph* 23), *Language* 24, No. 2.
- Karelitz, S., R.L. Karelitz, and L.S. Rosenfeld
1959 "Infants' Vocalizations and Their Significance", in *Mental Retardation, Proceedings of the 1st International Congress on Mental Retardation* (New York, Grune & Stratton), pp. 439-446. [A correlated phonograph recording of the primary data of this study is published separately!].
- Karelitz, S. and V.R. Fisichelli
1962 "The Cry Thresholds of Normal Infants and Those with Brain Damage", *Journal of Pediatrics* 61:679.
- Kersta, L.G.
1962 "Voiceprint Identification", *Nature* 196:1253-57, and *Journal of the Acoustical Society of America* 34:725 (A).
1969 "Automated Acoustic-Signature Verification System", Voiceprint Laboratories (New Jersey).
- Kersta, L.G. and J.A. Colangelo
1969 "The Spectrographic Speech Patterns of Identical Twins", Voiceprint Laboratories (New Jersey).
- Koenig, W., H.K. Dunn, and L.Y. Lacy
1946 "The Sound Spectrograph", in *Technical Aspects of Visible Speech* (= *Bell System Monograph B-1415*) and *Journal of the Acoustical Society of America* 17:19-49.
- Kopp, G.A. and H.C. Green
1946 "Basic Phonetic Principles of Visible Speech", in *Technical Aspects of Visible Speech* (= *Bell System Monograph B-1415*) and *Journal of the Acoustical Society of America* 17: 74-89.

- 1944 (see Gray).
- Lacerda, A. de
1933 "Der Polychromograph", in *Proceedings of the 1st International Congress of Phonetic Sciences*, Amsterdam 1932, *Archives Néerlandaises de Phonétique Expérimentale VIII-IX*: 265-270.
1934 "Die Chromographie", *Archives Néerlandaises de Phonétique Expérimentale X*:65-109.
- Ladefoged, P. and R. Vanderslice
1967 "The 'Voiceprint' Mystique", *Working Papers in Phonetics* (UCLA).
- Liddell, M.H.
1924, 25, 27 *The Physical Characteristics of Speech Sound* (= *Purdue University Bulletins* 16, 23, and 28).
1940 *The Elements of Sound and Their Relation to Language*, (Urbana, University of Illinois Press), 136 pp.
- Lilly, J.C., A.M. Miller, and H.M. Truby
1967 "Reprogramming of the Sonic Output of the Dolphin: Sonic-Burst-Count Matching", *Communication Research Institute Scientific Report No. 0267* (Coconut Grove).
- Lilly, J.C., H.M. Truby, A.M. Miller, and F. Grissman
1967 "Acoustic Implications of Interspecies Communication", *Journal of Acoustical Society of America* 42:1164(A) [Paper I 10, ASA Meeting No. 74, Miami Beach].
- Lynip, A.W.
1951 "The Use of Magnetic Devices in the Collection and Analysis of the Preverbal Utterances of an Infant", *Genetic Psychology Monographs* 44:221-262.
- Montagu, A.
1962 "The History of the Dolphin", in *The Dolphin in History* (Los Angeles, Clark Memorial Library), pp. 3-30.
- Panconcelli-Calzia, G.
1941 *Dreitausend Jahre Phonetik* (Hamburg, Hansischer Gilden).
- Parmelee, A.H., Jr.
1962 "Infant Crying and Neurologic Diagnosis", *The Journal of Pediatrics* 61.5:801-802 (St. Louis).
- Pierce, J.R.
1969 "Whither Speech Recognition?", *Journal of the Acoustical Society of America* 46:1049-51.
- Potter, R.K.
1945 "Visible Patterns of Sound", *Science* 102:463-470.
1946 "Introduction to Technical Discussions of Sound Portrayal", in *Technical Aspects of Visible Speech* (= *Bell System Monograph B-1415*) and *Journal of the Acoustical Society of America* 17:1-3.
- Potter, R.K., G.A. Kopp, and H.C. Green
1947 *Visible Speech* (New York, D. Van Nostrand).
- Riesz, R.R. and L. Schott
1946 "Visible Speech Cathode-Ray Translator", in *Technical Aspects of Visible Speech* (= *Bell Telephone System Monograph B-1415*) and *Journal of the Acoustical Society of America* 17:50-61.
- Sapir, E.
1927 "Speech as a Personality Trait", *American Journal of Sociology* 32:892-905.
- Scripture, E.W.
1902 *The Elements of Experimental Phonetics* (New York, Scribner's, London, Arnold). [And see footnote 11 herewith!].
1933 "Analysis and Interpretation of Vowel Tracks", *Journal of the Acoustical Society of America* 5:148.
- Steinberg, J.C.
1934 "Application of Sound Measuring Instruments to the Study of Phonetic Problems", *Journal of the Acoustical Society of America* 6:16-24.
- Steinberg, J.C. and N.R. French
1946 "The Portrayal of Visible Speech", in *Technical Aspects of Visible Speech* (= *Bell Telephone System Monograph B-1415*, and *Journal of the Acoustical Society of America* 17:4-18.

- Stevens, K.N.
1971 "Sources of Inter- and Intra-Speaker Variability in the Acoustic Properties of Speech Sounds", *Proceedings of the VIIth International Congress of Phonetic Sciences* (Montreal), to be found on pp. 206-232 of this volume.
- Tosi, O., H. Oyer, W. Lashbrook, C. Pedrey, J. Nicol, and E. Nash
1971 "An Experiment on Voice Identification" (Michigan State University).
- Truby, H.M.
1958a "A Note on Visible and Indivisible Speech", in *Proceedings of the VIIIth International Congress of Linguists*, Oslo.
1958b "The Inseparability of Clustered Consonants", a 35-mm cineradiographic film (accompanying VIIIth International Congress of Linguists paper).
1958c *Correlation of Cineradiographic and Visual-Acoustic Analyses of Speech*, 16- or 35-mm copy of 35-mm original X-ray sound film and complete libretto comprising texts and commentary (= *Communication Research Institute Scientific Report* 0159) (Stockholm).
1959 *Acoustico-Cineradiographic Analysis Considerations* (= *Acta Radiologica Supplementum* 182) 227 pp., 1100 speechsound spectrograms, 318 X-ray cineframe plates (Stockholm).
1960 "Some Aspects of Acoustical and Cineradiographic Analysis of Newborn-Infant and Adult Phonation and Associated Vocal-Tract Activity", *Journal of the Acoustical Society of America* 31:1518(A). [Paper Q 12, ASA Meeting No. 60, San Francisco].
1962a "A Preliminary Report on Newborn-Infant Cry-Sound Analysis", *Communication Research Institute Scientific Report* 0362 (Stockholm).
1962b Review of Ernst Pulgram, *Introduction to the Spectrography of Speech*, in *Word* 18:376-387.
1963 "Segmentation and Mechanical Commutation", *IBM Research Journal* 298, Paper X 10-ASA Meeting No. 66 (Ann Arbor) also *Journal of the Acoustical Society of America* 35:1912(A).
1964a "Pleniphonetic Transcription in Phonetic Analysis", *Proceedings of IXth International Congress of Linguists*, Cambridge, Mass., 1962 (The Hague, Mouton).
1964b "The Homoneme", *IBM Research Journal* 312 (San Jose) and in *Word* 22:2-3.
1964c "Speech Synthesis: The Relevance of Pseudospeech to Linguistics", paper delivered before the Linguistic Circle of Stanford.
1965a "Duration as an Alternate Synthesis-Parameter for Intensity and Vowel-Quality", in *Proceedings of the Vth International Congress of Phonetic Sciences*, Münster, 1964 (Basel, Karger), pp. 551-555.
1965b "Matching Speech Cineradiography with Pseudospeech", with a high-speed 35-mm X-ray film with dubbed synthesized-speech sound-track, *Journal of the Acoustical Society of America* 37:1187(A) (San Jose). [Paper 19, ASA Meeting No. 69, Washington].
1965c "Evaluation of Para- and Neo-natal Infant Cry Analyses", *Communication Research Institute Scientific Report* 0165, National Institute of Child Health and Human Development project description (San Jose).
1965d "A Phonetic Macro-matrix for Pseudospeech Synthesis", Xth Annual National Conference on Linguistics, Linguistic Circle of New York.
1966a Review of Ilse Lehiste, *Some Acoustic Characteristics of Dysarthric Speech*, *Phonetica* 15: 243-246.
1966b "Neonatal Cry-sound Analysis in Pediatric Diagnosis", National Institute of Child Health and Human Development Research Grant 01985 (= *Communication Research Institute Scientific Report* 0266) (Coconut Grove).
1967a "The Detection of Brain Damage in Infants in Terms of Acoustigraphic Cry-sound Analysis Criteria", *Communication Research Institute Scientific Report* 1567 (United Cerebral Palsy Research and Educational Foundation Grant Application).
in press "Synthesized — i.e., Artificial — Speech, i.e., Pseudospeech", address prepared for delivery before the American Philosophical Society, 1967.
1967b "Language and Dolphins", in *Proceedings of the Xth International Congress of Linguists*, (Bucharest, 1967) (Bucharest, Publication of the Academy of the Socialist Republic of Rumania, 1970); also *Communication Research Institute Scientific Report* 0467 (Coconut Grove).
1967c "Phonetic, Acoustic-Phonetic, and Linguistic Aspects of Neurophysiological Research on

- Large Brains", *Communication Research Institute Scientific Report* 1967 [for NINDB Grant "Neurophysiological Research on Large Brains", J.C. Lilly, P.J. Morgane, H.M. Truby].
- 1967d "The Perception and Common Misperception of Infant Pre-speech", in *Proceedings of the VIIth International Congress of Phonetic Sciences*, Prague, 1967 (Academia, Prague, 1970); also *Communication Research Institute Scientific Report* 0567.
- 1967e "A Definition of Speech-sound Analysis, 'Speech Synthesis', and Speech", in *Linguistic Studies, Vol. 1: General Linguistics* (= *Word* 23); also *Communication Research Institute Scientific Report* 0667.
- 1967f "Pseudospeech Synthesis and the Phonette", *Journal of the Acoustical Society of America* 42:1163(A); and *Communication Research Institute Scientific Report* 2067 (Coconut Grove). [Paper I 5, ASA Meeting, No. 74, Miami Beach].
- 1967g "A Chronology of Related Newborn-Infant Cry and Cry-sound Activities", American Speech and Hearing Association Convention 43 Scientific Exhibit (Chicago); *Communication Research Institute Scientific Report* 2467 (Coconut Grove).
- 1967h "Cryprint Categories in Newborn-Infant Cry-sound Analysis", *Journal of the Acoustical Society of America* 42:1164(A); also, *Communication Research Institute Scientific Report* 2267 (Coconut Grove). [Paper I 11, ASA Meeting No. 74, Miami Beach].
- 1967i "Pre-natal Speech: A Speculation", *Communication Research Institute Scientific Report* 2567 (Coconut Grove).
- 1968 "Phoneme, Allophoneme, Phone, and Allophone", XIII Annual National Conference on Linguistics, Linguistic Circle of New York; *Communication Research Institute Scientific Report* 0268 (Coconut Grove).
- 1969 "The Phoneme-Grapheme Myth", SouthEastern Conference on Linguistics II (Gainesville), *Language and Linguistics Research Study* 4 (Coral Gables).
- 1970a "The Validity of 'Voiceprinting', 'Cryprinting', and Other 'Soundprinting'", *Language and Linguistics Research Study* 9 (Coral Gables).
- 1970b "Dolphins and Language", *Muse News* 2.6 (Miami Museum of Science); also *Language and Linguistics Research Study* 10 (Coral Gables).
- 1971a "Glossary of Selected Linguistic Terminology", in *Phone Discrimination Tests in Test of Phonic Skills*, K.J. Smith and H.M. Truby (New York, Harper & Row), also *Language and Linguistics Research Study* 11.
- 1971b "The Phoneme Illusion Invalidates Phonetic Transcription", *Language and Linguistics Study* No. 12(A).
- 1971c "A Note on Synthetic Speech — Term and Concept", *Language and Linguistics Study* 17.
- Truby, H.M., J.F. Bosma, and J. Lind
1965 *Newborn Infant Cry* (= *Acta Paediatrica Scandinavica Supplementum* 163), 166 crysound spectrograms, 228 X-ray cineframe tracings of projected 35-mm negative (Stockholm), 128 pp., with correlated phonograph recording of visible-acoustic data.
- Truby, H.M., J. Bosma, J. Lind, and P. Karlberg
1960 "Infant Cry Sounds: A Visual-Acoustic Analysis Technique" [An accompanying phonograph record demonstrates all visible-acoustic data]. *Communication Research Institute Scientific Report* 0260 (Stockholm). [65 pp., 213 crysound spectrograms].
- Truby, H.M. and J.C. Lilly
1967 "Psychoacoustic Implications of Inter-species Communication", *Journal of the Acoustical Society of America* 42:1181(A); also *Communication Research Institute Scientific Report* 2367 (Coconut Grove). [Paper S3, ASA Meeting No. 74, Miami Beach].
- Truby, H.M., J. Lind, and J.F. Bosma
1961 "Cry Sounds of the Newborn Infant", 81 crysound spectrograms, *Communication Research Institute Scientific Report* 0161 (Stockholm), with correlated phonograph recording of visible-acoustic data.
- Truby, H.M. and C. Wegelius
1958 "A Cineradiographic Study of Certain Speech Articulation Processes", *Bulletin Final — relations, communications et discussions, IIIe Congrès International de Radiophotographie (Photofluorography)*, (Stockholm). The discussion of this report is on pages 1281-1287.