Apr 1st, 8:00 AM

A Satellite Association Procedure

Anthony G. Lubowe
Analytical Mechanics Department, Bell Telephone Laboratories, Incorporated, Whippany, New Jersey 07981

Scholarly Commons Citation
Lubowe, Anthony G., "A Satellite Association Procedure" (1969). *The Space Congress® Proceedings*. 3.
https://commons.erau.edu/space-congress-proceedings/proceedings-1969-6th-v1/session-15/3

A SATELLITE ASSOCIATION PROCEDURE

Anthony G. Lubowe

Analytical Mechanics Department
Bell Telephone Laboratories, Incorporated
Whippany, New Jersey 07981

SUMMARY

A procedure is derived for estimating the consistency of a radar observation of an object with a prediction of an orbiting object. This procedure may be of use as an association procedure, i.e., to insure that data is indeed being taken on an intended object before "associating" the new data with old data on the object.

Specifically we show that, with reasonable assumptions about the observational and prediction errors, a quadratic form associated with the position error vector has a chi-square distribution with 3 degrees of freedom. Thus we can compute the probability of the residual if the observation and the prediction come from the same object. A low probability is taken as an indication that the prediction and observation refer to different objects.

The computational procedure is described in detail, and a Monte Carlo run is included to demonstrate the correctness of the procedure.

INTRODUCTION

By a satellite association procedure we mean a test or series of tests to decide whether an observation of a satellite is "consistent" with a prediction of the same satellite. The test is used to determine whether or not to associate the observation (and succeeding ones) with the data stored for the predicted satellite. "Consistency" is taken to mean that we believe the observation has been taken on the same object that we have been predicting, or alternately, that we are willing to accept the possibility that we are falsely matching two different objects.

AN ASSOCIATION TEST

We have a predicted satellite which we can denote by a geocentric radius vector $\vec{r}_p$ and a velocity vector, $\vec{v}_p$, which we place at the satellite. We assume that the components of $\vec{r}_p$ in an equatorial coordinate system, say x_p, y_p, z_p are sample values from a trivariate normal distribution with covariance matrix A_p, which is of course 3×3.*

We also have an observed satellite which we can denote by a geocentric radius vector $\vec{r}_o$ with components x_o, y_o, z_o. Actually, we observe some components, say, R,A,E of a topocentric vector $\vec{r}_T$; then $\vec{r}_o = \vec{r}_T + \vec{R}_s$ where $\vec{R}_s$ is the nonrandom geocentric radius vector of the observing site. We can assume that x_o, y_o, z_o are sample values from a trivariate normal distribution with covariance matrix A_o, again 3×3.

We define a residual vector $\vec{\Delta} = \vec{r}_o - \vec{r}_p$. It can be shown that the components of this vector are sample values of a trivariate normal distribution with covariance matrix $A = A_p + A_o$. (See Appendix A, Theorem 1.) Furthermore, the quadratic form

$$\varphi = \vec{\Delta}^T A^{-1} \vec{\Delta}$$

is a sample value of a random variable which has a chi-square distribution with three degrees of freedom (Ref. 1, pp. 348-349). The value of φ provides a single, meaningful test of the likelihood that the observed and predicted satellites are the same.

For example,*

Probability of $\varphi < 12.8 \approx 0.995$.

This is comparable to a "3-sigma" value when dealing with a normal distribution.

This is intuitive rather than precise language. Actually, we should say something like the following: If the residual vector $\vec{\Delta}$ is indeed chosen from the described distribution (i.e., if $\vec{r}_p$ and $\vec{r}_o$ refer to the same satellite, and differ only because they are sample values of random variables with the described distributions), then the probability that $\varphi < 12.8$ is 0.995. Thus a value of $\varphi > 12.8$ is rather unlikely (with probability of only 0.005) and we may choose to take it as an indication that $\vec{\Delta}$ is not a sample value of the described distribution, i.e., that $\vec{r}_p$ and $\vec{r}_o$ refer to different satellites. The corresponding situation with a normal distribution is as follows: A sample is chosen, possibly from a normal

---

* $\vec{v}_p$ is required only because we wish to consider an orbital coordinate system which has an axis in the $\vec{v}_p$ direction. We obtain $\vec{r}_p$ and $\vec{v}_p$ at any desired time, say t, the time of an observation, from orbital elements.

* See Table II. Since $P(\varphi > 12.838) = .005$, $P(\varphi < 12.8) \approx .995$.

distribution with mean zero and variance σ^2. If the sample was actually chosen from the normal distribution, the probability of a sample value with magnitude $\leq 3\sigma$ is 0.997. Thus a sample value $> 3\sigma$ is rather unlikely (with probability of only 0.003) and we may choose to take it as an indication that the sample was not chosen from the specified normal distribution.

We consider the computations required to implement this test in the next section.

COMPUTATIONAL CONSIDERATIONS

To compute the quadratic form φ we must obtain the residual vector $\vec{\Delta}$ and the covariance matrix A. We proceed as follows. We have observed an object and obtained values of R, A, E referred to our sensor-based coordinate system. We assume that these are sample values from a trivariate normal distribution with covariance matrix

$$A_1 = \begin{bmatrix} \sigma_R^2 & 0 & 0 \\ 0 & \sigma_A^2 & 0 \\ 0 & 0 & \sigma_E^2 \end{bmatrix} \qquad (1)$$

where the σ's are to be based upon the sensor accuracy. If one feels that the sensor measurements are correlated this is easily handled by non-zero off-diagonal elements in (1).

We convert R, A, E to rectangular coordinates, that is, we obtain the components of the vector called $\vec{r}_T$ in the previous section. The components of $\vec{r}_T$ are normally distributed random variables (by Theorem 2 of Appendix A) if $\vec{r}_T$ and [R,A,E] are connected by the linear transformation

$$\vec{r}_T = C_1 [R,A,E]^T \qquad (2)$$

and the covariance matrix of $\vec{r}_T$ is

$$A_2 = C_1 A_1 C_1^T \qquad (3)^*$$

We assume that (3) is indeed true, that is we neglect nonlinearity in the computation of the new covariance matrix, but we of course use the exact transformation instead of (2) to convert the means. The nonlinearities should be negligible for deviations from the means of the order of several standard deviations. We consider the magnitude of these nonlinearities as well as giving explicit expressions for the transformation and the matrix C_1 in Appendix B.

The conversion from $\vec{r}_T$ to $\vec{r}_o$ (the geocentric radius vector to the observation in an inertial ractangular equatorial coordinate system) is linear as required by Theorem 2 of Appendix A. Thus

$$\vec{r}_o = C_2 \vec{r}_T + b \qquad (4)$$

and

$$A_o = C_2 A_2 C_2^T. \qquad (5)$$

(See Appendix B for the explicit form of C_2.) Finally, we obtain, from (5) and (3)

$$A_o = (C_2 C_1) A_1 (C_2 C_1)^T \qquad (6)$$

We have thus obtained $\vec{r}_o$ and A_o; now we must compute $\vec{r}_p$ and A_p.

We obtain $\vec{r}_p$ at the time of the observation from a set of osculating orbital elements valid at, or near to, the time of the observation. The procedure for converting standard osculating orbital elements valid at some time, t, to position and velocity components at the same time is well known.

Given $\vec{r}_p = [x_p, y_p, z_p]^T$ we can compute the residual vector $\vec{\Delta} = \vec{r}_o - \vec{r}_p$. Also, $\vec{r}_p$ and $\vec{v}_p$ (the corresponding predicted velocity) determine the orbital coordinate system,* so that

$$\begin{bmatrix} \Delta H + r_p \\ \Delta u \\ \Delta\beta \end{bmatrix} = C \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} = C \vec{r}_o \qquad (7)$$

where the elements of the 3×3 matrix C (which are functions of $\vec{r}_p$ and $\vec{v}_p$) are given in Appendix B. Actually we need the inverse of (7), i.e.

$$\vec{r}_o = \begin{bmatrix} x_o \\ y_o \\ z_o \end{bmatrix} = C_3 \begin{bmatrix} r_p + \Delta H \\ \Delta u \\ \Delta\beta \end{bmatrix} \qquad (8)$$

---

[*] If the measurements are actually phased array radar coordinates (R, sin α, sin β), and are normally distributed but not necessarily statistically independent, then it can be shown[2] that we can replace A_2 by the appropriate matrix and proceed.

[*] Specifically, we define a system with axes along the geocentric radius vector of the predicted satellite (H), along the velocity vector (u), and along the normal to the orbit plane (β). Note that the β axis is normal to both the H and u axes (which determine the orbit plane); however the H and u axes are not orthogonal except in the special case of circular orbits.

where the elements of $C_3 = C^{-1}$ are also given in Appendix B.* (It can be verified that $C\vec{r}_p = [r_p,0,0]^T$ and $\vec{r}_p = C_3[r_p,0,0]^T$ as they must for (7) and (8) to be valid when $\vec{r}_p = \vec{r}_o$.)

Since (8) represents a transformation of the class described by Theorem 2 of Appendix A, we have

$$A_p = C_3 \, A_3 \, C_3^T \tag{9}$$

where we take

$$A_3 = \begin{bmatrix} \sigma_H^2 & 0 & 0 \\ 0 & \sigma_u^2 & 0 \\ 0 & 0 & \sigma_\beta^2 \end{bmatrix} \tag{10}$$

with the σ's to be based upon the prediction accuracy. If, as is frequently the case, the orbit determination procedure produces the covariance matrix as well as an estimate of the elements themselves, this could be used simply to compute A_3, which would then in general be nondiagonal. Since we have A_o and A_p from Equations (6) and (9) we can compute $A = A_o + A_p$.

Finally we can compute

$$\varphi = \vec{\Delta}^T \, A^{-1} \, \vec{\Delta} \tag{11}$$

by performing the indicated inversion and multiplications. Explicit expressions could be used for any or all of Equations (6), (9), or (11) if possible computational simplifications were especially necessary.

NUMERICAL RESULTS

A Monte Carlo run was made to test some of our assumptions and calculations. (Specifically, we wish to check the analytic proofs in Appendix A, the coordinate transformations in Appendix B, the computer programs used, as well as the allowability of the neglect of any nonlinearity in the transformations. The tracker and nominal satellite characteristics are given in Table I. The statistics assumed for the observational and prediction errors are also given. Triples of observation errors (ΔR,ΔA,ΔE) and prediction errors (ΔH,Δu,Δβ) were drawn from normal distributions with zero means and standard deviations as given in Table I. These were used to modify the nominal orbit to produce a simulated observation and a simulated prediction. The association test was applied to this pair, i.e., a value of φ was computed. The results for 10,000 trials are given in Table II as $N_i\left(\varphi > \chi_i^2\right)$ i.e., the number of computed values of φ actually found to be greater than χ_i^2. We also tabulate the theoretical values of the cumulative density function, P, for a chi-square distribution with three degrees of freedom. The values of P_i (times the number of realizations n = 10,000) should be compared with the actually obtained values, N_i. The agreement is seen to be good.

In passing we note that the Monte Carlo run required 0.0808 hours of GE-635 computer time for the 10,000 samples or less than 30 milliseconds for one evaluation of the association test. This includes the time spent in random number generation and production of the simulated observation and prediction and is thus a pessimistic upper limit to the time actually required in practice for one evaluation of the association test.

The results are further analyzed in Table III where the empirical results (the 10,000 values of φ) are divided into 51 groups, i.e., $N_i(\chi_{i-1}^2 < \varphi < \chi_i^2)$ i = 1,...,51. These are to be compared with the theoretical results, i.e., $nP_i(\chi_{i-1}^2 < \varphi < \chi_i^2)$, i = 1,...,51, which are also given in Table III.

It can be shown (see, for instance, Ref. 2, pp. 299-305 and Ref. 3, pp. 9-12 for some discussion) that

$$Q_{50} = \sum_{i=1}^{51} \frac{(N_i - nP_i)^2}{nP_i}$$

has a limiting distribution (as n → ∞) which is chi-square with 50 degrees of freedom. This random variable can be used to evaluate the goodness-of-fit of our empirical results to the theoretical results from a chi-square distribution (ν = 3).

Using Table III we compute

$$Q_{50} = 43.2.$$

Interpolating in Table 7 of Ref. 3 for ν = 50 we find

$$P(Q_{50} > 43.2) = 0.74$$

so that a value of Q_{50} as large as obtained from our result might have arisen through random sampling fluctuations almost 3 out of 4 times. (A similar result, $Q_{10} = 7.4$, was obtained testing with 11 intervals. Since $P(Q_{10} > 7.4) = 0.69$, the effect of the arbitrary choice of interval does not appear too great in this case.)

Thus, there is no reason to reject the hypothesis that the distribution of φ is chi-square with 3 degrees of freedom, and the desired check on our analysis has been obtained. Therefore the value of φ can be used as a single, statistically consistent test to decide whether an observation has indeed been taken of an object for which we have a prediction, or whether the observation and the prediction refer to two different objects.

---

* $C^{-1} \neq C^T$ unless the orbit is circular, in which case the transformation given by (7) is orthogonal.

Acknowledgments

I am grateful to G. W. Grube, Mrs. M. S. Katz, L. K. Russell, and Mrs. C. B. Wood for programming assistance. I also appreciate comments from A. J. Claus, Mrs. J. Eilbott, F. T. Geyling, J. C. Hemmer, R. Levine, R. A. Marth, A. J. Pasqua, C. S. Sherrerd, and D. D. Slavinskas.

These studies were performed under Department of Defense Contract No. DA-30-069-AMC-333(Y).

References

1. R. V. Hogg and A. T. Craig, Introduction to Mathematical Statistics, 2nd edition (New York: MacMillan, 1965).
2. D. D. Slavinskas, "Statistics of Conversion from Phased Array Radar Coordinates to a Cartesian Coordinate System", December 23, 1966, unpublished work.
3. E. S. Pearson and H. O. Hartley, Editors, Biometrika Tables for Statisticians, Vol. 1, 3rd Edition (Cambridge University Press, 1966).

Table I: Monte Carlo Run (n = 10,000)

$a = 2.125\times10^{7}$ ft
$i = 45°$
$\Omega = 0$

$\alpha = 90°$
$R_E = 2.0925738\times10^{7}$ ft

$\sigma_R = 250$ ft
$\sigma_A = \sigma_E = 8.7266\times10^{-4}$

$\sigma_H = 4,400$ ft
$\sigma_u = 440,000$ ft
$\sigma_\beta = 44,000$ ft

$e = 0$
$\omega = 0$
$\tau = -1.29691559\times10^{3}$ sec

$\delta = \delta'' = 45°$
$k = 1.407645\times10^{16}$ ft^3/sec^2

Table II: $P_i\left(\varphi > \chi_i^2\right)$[*] and $N_i\left(\varphi > \chi_i^2\right)$

χ_i^2	P_i	N_i	χ_i^2	P_i	N_i	χ_i^2	P_i	N_i
.0	1.0000	10000	1.9	.5934	5896	6.251	.1000	983
.0717	.9950	9956	2.0	.5724	5702	6.4	.0937	920
.1	.9918	9925	2.2	.5320	5307	6.6	.0858	851
.115	.9900	9901	2.366	.5000	4957	6.8	.0786	781
.2	.9776	9769	2.4	.4936	4898	7.0	.0719	716
.216	.9750	9747	2.6	.4575	4498	7.2	.0658	651
.3	.9600	9584	2.8	.4235	4178	7.4	.0601	597
.352	.9500	9483	3.0	.3916	3830	7.6	.0550	539
.4	.9402	9382	3.2	.3618	3452	7.8	.0503	502
.5	.9189	9181	3.4	.3340	3250	7.815	.0500	498
.584	.9000	8983	3.6	.3080	3006	8.0	.0460	466
.6	.8964	8951	3.8	.2837	2785	8.2	.0421	421
.7	.8732	8706	4.0	.2615	2564	8.4	.0384	390
.8	.8495	8457	4.108	.2500	2461	8.6	.0351	348
.9	.8254	8209	4.2	.2407	2372	8.8	.0321	317
1.0	.8013	7966	4.4	.2214	2196	9.0	.0293	287
1.1	.7771	7705	4.6	.2035	2024	9.2	.0268	266
1.2	.7530	7488	4.8	.1870	1856	9.348	.0250	252
1.213	.7500	7449	5.0	.1718	1691	9.4	.0244	247
1.3	.7291	7261	5.2	.1577	1560	9.6	.0223	221
1.4	.7055	7034	5.4	.1447	1426	9.8	.0203	205
1.5	.6823	6793	5.6	.1328	1295	10.0	.0186	186
1.6	.6594	6597	5.8	.1217	1186	11.345	.0100	96
1.7	.6369	6368	6.0	.1116	1099	12.838	.0050	44
1.8	.6149	6123	6.2	.1023	1001	16.266	.0010	7

---

[*] Based upon Tables 7 and 8 of Ref. 3.

TABLE III*: $N_i\left(\chi^2_{i-1} < \varphi < \chi^2_i\right)$ and $nP_i\left(\chi^2_{i-1} < \varphi < \chi^2_i\right)$

χ^2_i	N_i	nP_i	χ^2_i	N_i	nP_i
.2	231	224	5.2	131	141
.4	387	374	5.4	134	130
.6	431	438	5.6	131	119
.8	494	469	5.8	109	111
1.0	491	482	6.0	87	101
1.2	478	483	6.2	98	93
1.4	454	475	6.4	81	86
1.6	437	461	6.6	69	79
1.8	474	445	6.8	70	72
2.0	421	425	7.0	65	67
2.2	395	404	7.2	65	61
2.4	409	384	7.4	54	57
2.6	400	361	7.6	58	51
2.8	320	340	7.8	37	47
3.0	348	319	8.0	36	43
3.2	288	298	8.2	45	39
3.4	292	278	8.4	31	37
3.6	244	260	8.6	42	33
3.8	221	241	8.8	31	30
4.0	221	224	9.0	30	28
4.2	192	208	9.2	21	25
4.4	176	193	9.4	19	24
4.6	172	179	9.6	26	21
4.8	168	165	9.8	16	20
5.0	165	152	10.0	19	17
			∞	186	186

---

* $\chi_o = 0$, $n = 10000$.

APPENDIX A

Two Theorems on the Combination of Normally Distributed Vectors

Theorem 1:[*]

If $X_{(1)}, X_{(2)}, \ldots, X_{(m)}$ are m random vectors whose m sets of n components are mutually stochastically independent random variables from multivariate normal distributions with matrices $\mu_i (i = 1, \ldots, m)$ of means and positive definite covariance matrices $A_i (i = 1, \ldots, m)$, then the linear combination

$$Z = c_1X_{(1)} + c_2X_{(2)} + \ldots + c_mX_{(m)},$$

where the c_i are real numbers and not all zero, has a multivariate normal distribution with matrix

$$c_1\mu_1 + c_2\mu_2 + \ldots + c_m\mu_m$$

of means and positive definite covariance matrix

$$c_1^2A_1 + c_2^2A_2 + \ldots + c_m^2A_m.$$

Proof:

We give a proof for the specific case m = 2, n = 3. For short, we define $X_{(1)} = X$, $X_{(2)} = Y$, $c_1 = c$, $c_2 = d$, $A_1 = A$, $A_2 = B$.

Let X_1, X_2, X_3 have a trivariate normal distribution with matrix μ of means and positive definite covariance matrix A. If we let $X^T = [X_1, X_2, X_3]$, then the moment-generating function $M(t_1, t_2, t_3)$ of this joint distribution of probability is[**]

$$E\left(e^{t^Tx}\right) = \exp\left(t^T\mu + \frac{t^TAt}{2}\right) \tag{A1}$$

where $t^T = [t_1, t_2, t_3]$, and we have used superscript T for "transpose".

Similarly, let Y_1, Y_2, Y_3 have a trivariate normal distribution with matrix ν of means and positive definite covariance matrix B. If we let $Y^T = [Y_1, Y_2, Y_3]$, then the moment-generating function $N(t_1, t_2, t_3)$ of this joint distribution of probability is

$$E\left(e^{t^Ty}\right) = \exp\left(t^T\nu + \frac{t^TBt}{2}\right) \tag{A2}$$

Consider the linear function Z of $X_1, X_2, X_3, Y_1, Y_2, Y_3$ which is defined by $Z = cX + dY$ where c and d are real numbers and not both zero. The moment-generating function $G(t_1, t_2, t_3)$ of the distribution of Z is given by

$$\begin{aligned} E\left(e^{t^Tz}\right) &= E\left(e^{t^T(cx+dy)}\right) \\ &= E\left(e^{ct^Tx}\right) E\left(e^{dt^Ty}\right) \end{aligned} \tag{A3}$$

if X and Y are mutually stochastically independent.

Thus, replacing t^T in (A1) and (A2) by ct^T and dt^T, respectively, we obtain from (A3)

$$\begin{aligned} G(t^T) &= \exp\left(ct^T\mu + \frac{t^TAtc^2}{2}\right) \\ &\cdot \exp\left(dt^T\nu + \frac{t^TBtd^2}{2}\right) \\ &= \exp\left\{t^T(c\mu+d\nu) + \frac{t^T(c^2A+d^2B)t}{2}\right\} \end{aligned} \tag{A4}$$

Equation (A4) is the moment-generating function of a multivariate normal distribution with matrix $(c\mu+d\nu)$ of means and positive definite covariance matrix (c^2A+d^2B).

The extension to multivariate normal distributions, i.e., $X^T_{(1)} = [X_{11}, X_{12}, \ldots, X_{1n}]$, $X^T_{(2)} = [X_{21}, X_{22}, \ldots, X_{2n}], \ldots$, $X^T_{(n)} = [X_{m1}, X_{m2}, \ldots, X_{mn}]$, $Z = c_1X_{(1)} + \ldots + c_mX_{(m)}$, is immediate, requiring only notational changes.

For the case discussed in the text we take $X = \vec{r}_o$, $Y = \vec{r}_p$, $A = A_o$, $B = A_p$. Then, $c = 1$, $d = -1$, $c^2 = d^2 = 1$, and the stated result follows from our theorem.

The second theorem is a variation of the first but it does not seem obvious how to write a single theorem which is equivalent to both.

Theorem 2:

If X is a random vector whose n components $X_1, X_2, \ldots, X_n$ are normally distributed random variables with matrix μ of means and positive definite covariance matrix A, then the linear combination

$$Z = CX + b$$

---

[*] This Theorem is a generalization of Ref. 2, p. 347, Example 1, and contains a well-known result. Although there must be many published versions I am not aware of any.

[**] See Ref 1, Chapter 13, particularly pp. 343-347.

(where C is an $n \times n$ matrix of real numbers not all zero and b is an $n \times 1$ matrix of real numbers) has a multivariate normal distribution with matrix $C\mu + b$ of means and positive definite covariance matrix CAC^T.

Proof:

The moment-generating function $M(t_1,t_2,\ldots,t_n)$ of X is

$$E\left(e^{t^T x}\right) = \exp\left(t^T\mu + \frac{t^T A t}{2}\right)$$

where $t^T = [t_1,t_2,\ldots,t_n]$.

We introduce the change of variable

$$t^T x = s^T(z-b) = s^T C x$$

into (A5) obtaining

$$E\left(e^{s^T z}\right) e^{-s^T b} = \exp\left(s^T C\mu + \frac{s^T CAC^T s}{2}\right)$$

or

$$E\left(e^{s^T z}\right) = \exp\left(s^T[C\mu+b] + \frac{s^T[CAC^T]s}{2}\right) \quad \text{(A6)}$$

But (A6), i.e., the moment-generating function of Z, is the moment-generating function of a multivariate normal distribution with matrix $C\mu + b$ of means and positive definite covariance matrix CAC^T as was to be shown.

APPENDIX B

Coordinate Conversions

1. $\vec{r}_T$ from R,A,E:

If we define $\vec{r}_T = [x_T,y_T,z_T]^T$ where x_T points East, y_T North, and z_T up, then

$$\begin{aligned} x_T &= R \cos E \sin A \\ y_T &= R \cos E \cos A \\ z_T &= R \sin E \end{aligned} \quad \text{(B1)}$$

The linearized version is

$$\begin{pmatrix} \delta x \\ \delta y \\ \delta z \end{pmatrix} = \begin{pmatrix} \sin \hat{A} \cos \hat{E} & \hat{R} \cos \hat{A} \cos \hat{E} & -\hat{R} \sin \hat{A} \sin \hat{E} \\ \cos \hat{A} \cos \hat{E} & -\hat{R} \sin \hat{A} \cos \hat{E} & -\hat{R} \cos \hat{A} \sin \hat{E} \\ \sin \hat{E} & 0 & \hat{R} \cos \hat{E} \end{pmatrix} \cdot \begin{pmatrix} \delta R \\ \delta A \\ \delta E \end{pmatrix} \quad \text{(B2)}$$

The 3×3 matrix in (B2) is what we have called C_1. We have denoted nominal values, say of R, as $\hat{R}$, and deviations from nominal, say $R-\hat{R}$, as δR.

To see the order of the errors introduced by use of (B2), consider the following example. Let $\hat{R} = 1{,}000{,}000$ ft, $\hat{A} = 150°$, $\hat{E} = 45°$. Consider deviations from nominal of the order of 0.1% (10^{-3}). Then $\delta R = 1000$ ft, $\delta A = 0.150°$, $\delta E = 0.045°$. Using (B1) to compute the nominal $\vec{r}_T$ we obtain

$$\begin{aligned} \hat{x}_T &= 353{,}553.39 \\ \hat{y}_T &= -616{,}372.43 \\ \hat{z}_T &= 707{,}106.78 \end{aligned}$$

Using (B1) to compute $\vec{r}_T$ we obtain

$$\begin{aligned} x_{T1} &= 352{,}024.15 \\ y_{T1} &= -613{,}426.88 \\ z_{T1} &= 708{,}369.57 \end{aligned}$$

The values obtained using the linearized version (B2), are

$$\begin{aligned} x_{T2} &= \hat{x}_T + \delta x = 352{,}026.08 \\ y_{T2} &= \hat{y}_T + \delta y = -613{,}429.46 \\ z_{T2} &= \hat{z}_T + \delta z = 708{,}369.23 \end{aligned}$$

The discrepancies due to the linearization are thus

$$\begin{aligned} (x_{T2}-x_{T1})/x_{T1} &= 5.481 \times 10^{-6} \\ (y_{T2}-y_{T1})/y_{T1} &= 4.202 \times 10^{-6} \\ (z_{T2}-z_{T1})/z_{T1} &= -.474 \times 10^{-6}, \end{aligned}$$

or of the order of 10^{-6} as is to be expected, since second order effects ($10^{-3} \times 10^{-3}$) were neglected in the derivation of Equation (B2).

2. $\vec{r}_o = C_2\vec{r}_T + b$ (Equation (4)):

We describe the position of the tracker at the time of the observation by its right ascension, α, and its geocentric and geodetic latitudes, δ'' and δ. If we denote the earth's equatorial radius by R_E, then

$$C_2 = \begin{pmatrix} -\sin\alpha & -\cos\alpha \sin\delta & \cos\alpha \cos\delta \\ \cos\alpha & -\sin\alpha \sin\delta & \sin\alpha \cos\delta \\ 0 & \cos\delta & \sin\delta \end{pmatrix} \quad \text{(B3)}$$

and

$$b^T = (R_E \cos\alpha \cos\delta'', R_E \sin\alpha \cos\delta'', R_E \sin\delta'') \quad \text{(B4)}$$

3. Orbital Coordinates from $\vec{r}_o$ (Equation (7)):

We obtain the matrix C in two steps. First we note that we can transform from $\vec{r}_o$ to $\vec{X}$ where $\vec{X} = [X,Y,Z]^T$ (X along the radius vector, Z normal to the orbit plane, and Y chosen to produce a right-handed orthogonal system) using

$$\vec{X} = \bar{C}\,\vec{r}_o = \begin{bmatrix} \bar{\ell}_1 & \bar{m}_1 & \bar{n}_1 \\ \bar{\ell}_2 & \bar{m}_2 & \bar{n}_2 \\ \ell_3 & m_3 & n_3 \end{bmatrix} \vec{r}_o \qquad \text{(B5)}$$

where

$\bar{\ell}_1 = \cos\theta \cos\Omega - \sin\theta \cos i \sin\Omega$

$\bar{\ell}_2 = -\sin\theta \cos\Omega - \cos\theta \cos i \sin\Omega$

$\ell_3 = \sin i \sin\Omega$

$\bar{m}_1 = \cos\theta \sin\Omega + \sin\theta \cos i \cos\Omega$

$\bar{m}_2 = -\sin\theta \sin\Omega + \cos\theta \cos i \cos\Omega$

$m_3 = -\sin i \cos\Omega$

$\bar{n}_1 = \sin\theta \sin i$

$\bar{n}_2 = \cos\theta \sin i$

$n_3 = \cos i$

(and θ is the angle in the orbit plane from the x-y plane to the radius vector, Ω is the angle in the x-y plane from the x-axis to the orbit plane, and i is the angle between the x-y plane and the orbit plane).

Care must be taken when transforming from the rectangular system X, Y, Z to the oblique system ΔH, Δu, $\Delta\beta$. Let us first note that the cosines of the angles between the X and Y axes and the velocity vector (the Δu axis) are

$$\begin{aligned} \cos(X,\Delta u) &= \alpha\bar{\ell}_1 + \beta\bar{m}_1 + \gamma\bar{n}_1 \equiv C_X \\ \cos(Y,\Delta u) &= \alpha\bar{\ell}_2 + \beta\bar{m}_2 + \gamma\bar{n}_2 \equiv C_Y \end{aligned} \qquad \text{(B6)}$$

where $[\alpha,\beta,\gamma] = \frac{1}{V}\,[\dot{x},\dot{y},\dot{z}]$ and $V = (\dot{x}^2+\dot{y}^2+\dot{z}^2)^{1/2}$. Equations (B6) follow from the fact that $[\bar{\ell}_1,\bar{m}_1,\bar{n}_1]$, $[\bar{\ell}_2,\bar{m}_2,\bar{n}_2]$, $[\alpha,\beta,\gamma]$ are the direction cosines of the X, Y, and Δu axes in the x-y-z system.*

If we define unit vectors $\underline{e}_X$, $\underline{e}_Y$, $\underline{e}_Z$; $\underline{e}_H$, $\underline{e}_u$, $\underline{e}_\beta$ along the designated axes, then

$$\begin{aligned} \underline{e}_X &= \underline{e}_H \\ \underline{e}_Y &= (1/C_Y)\underline{e}_u - (C_X/C_Y)\underline{e}_H \\ \underline{e}_Z &= \underline{e}_\beta \end{aligned}$$

Therefore,

$$\begin{aligned} r_p + \Delta H &= X - (C_X/C_Y)Y \\ \Delta u &= (1/C_Y)Y \\ \Delta\beta &= Z \end{aligned} \qquad \text{(B7)}$$

Combining (B7) and (B5), we obtain

$$\begin{bmatrix} \Delta H + r_p \\ \Delta u \\ \Delta\beta \end{bmatrix} = C\vec{r}_o \qquad (7)$$

where

$$C = \begin{bmatrix} \bar{\ell}_1 - \frac{C_X}{C_Y}\bar{\ell}_2 & \bar{m}_1 - \frac{C_X}{C_Y}\bar{m}_2 & \bar{n}_1 - \frac{C_X}{C_Y}\bar{n}_2 \\ \bar{\ell}_2/C_Y & \bar{m}_2/C_Y & \bar{n}_2/C_Y \\ \ell_3 & m_3 & n_3 \end{bmatrix} \qquad \text{(B8)}$$

4. <u>$\vec{r}_o$ from Orbital Coordinates (Equation (8)):</u>

We obtain from (B5)*

$$\vec{r}_o = \bar{C}^T\,\vec{X} = \bar{C}^T[X,Y,Z]^T.$$

Since $X = (r_p+\Delta H) + C_X\Delta u$, $Y = C_Y\Delta u$, $Z = \Delta\beta$, we have

$$\vec{r}_o = C_3\begin{bmatrix} r_p + \Delta H \\ \Delta u \\ \Delta\beta \end{bmatrix} \qquad (8)$$

where

$$C_3 = \begin{bmatrix} \bar{\ell}_1 & C_X\bar{\ell}_1 + C_Y\bar{\ell}_2 & \ell_3 \\ \bar{m}_1 & C_X\bar{m}_1 + C_Y\bar{m}_2 & m_3 \\ \bar{n}_1 & C_X\bar{n}_1 + C_Y\bar{n}_2 & n_3 \end{bmatrix} \qquad \text{(B9)}$$

Using properties of $\bar{C}$, such as $\bar{\ell}_1^2 + \bar{\ell}_2^2 + \bar{\ell}_3^2 = 1$ and $\bar{\ell}_1\bar{\ell}_2 + \bar{m}_1\bar{m}_2 + \bar{n}_1\bar{n}_2 = 0$, one can verify that $CC_3 = C_3C = I$ as it must for both (7) and (8) to be correct.

---

* The relation, $\cos(Z,\Delta u) = \alpha\ell_3 + \beta m_3 + \gamma n_3 = 0$, is identically satisfied. Note that $C_X = 0$ and $C_Y = 1$ for circular orbits.

* Since $\bar{C}$ is orthogonal, $\bar{C}^{-1} = \bar{C}^T$

THE IMPACT OF SPEECH COMMUNICATION WITH COMPUTERS

Wayne A. Lea
Man-Computer Systems Branch
NASA Electronics Research Center
Cambridge, Massachusetts

Summary

Manned space flights and other aerospace applications are requiring increasingly sophisticated multifunctional computer facilities and more versatile man-computer interaction capabilities. Research has disclosed that, among the several possible modalities of man-computer communication, speech is particularly promising, especially for manned spaceflight. The many specific advantages of speech justify a comprehensive program in automatic speech recognition and synthesis, segmentation and production of natural-flowing continuous speech, and development of relevant physiological and linguistic models.

Aerospace-promoted research on speech communication with computers contributes much to improved understanding of human communication processes, phonological and syntactic models, and mechanisms of speech production and perception. Mechanisms for aiding the deaf, blind, and others physically handicapped are resulting from efforts to mechanically analyze or synthesize speech. Such studies also are suggesting improved methods for language learning, including visual displays of correct and student-produced pronunciation patterns. Such results suggest the widespread scientific and social impact of aerospace techniques for mechanical processing of speech.

1. Introduction

Space is man's new frontier. Like previous frontiers, it has posed many awesome practical problems that require increased stamina, new techniques, and a fresh understanding of man's relationship to his environment. Old problems also take on a new significance. Take for example the common cold and household diseases like influenza. Apollo 7 and 8 showed us how the common illnesses, when experienced in space, can be threatening, debilitating, and intricate influences on the success or failure of future manned space missions. News commentators and opinion makers in America commented on this at some length, suggesting it may well lead to a renewed and stepped-up attack on the common cold.

Keeping the man a working, productive part of the spaceborne system is a complex challenge. All his talents and special abilities--those capabilities that make manned space flight a profitable adventure--must be maintained at an optimal level of utility. Optimum use of the man, and the complex machinery he controls, particularly demands understanding of how the man can best interact with machines. Thus we have the popular new field of man-machine interaction. If man is to succeed in space, the capabilities and limitations of the complete man-machine system must be judiciously and iteratively reviewed and evaluated.

In this paper, attention will be drawn to a particularly versatile form of man-machine communication and to the impact of work in this field on the advancement of space science and the well-being of our earth-bound society. The particular form of man-machine communication of interest here is man's most natural, universally-used, fast-acting, and flexible communication modality: speech. The age-old problems of how man speaks and listens take on new significance in the man-machine interface.

Man has, for centuries, communicated by speech. However, with few exceptions, he has directed his spoken conversation to other men. There are the times when he talks to his dog, his cat, his mimicking bird, or other pet or animal. Other times he rages in anger at his "blinkety-blank hammer", his "stubborn and stupid" automobile, his spastic television set, or his other mechanisms. But now, with computers, comes the possibility of a mechanism's comprehension and reply to such spoken commands, curses, involuntary remarks, or other conversation.

Following the imaginative and glamorous (albeit, rarely fulfilled) promises of the proponents of artificial intelligence and versatile robots, one may visualize conversational computers with linguistic sophistication and some appearance of high intelligence. The idea of brilliant electronic brains with which one can communicate in natural, conversational English arouses a variety of emotions, including unbelief, awe, fear, fascination, and so forth. Yet, there is a practical, realistic side to such fascinating searches for vocally literate machines. Practical problems do exist where speech communication with computers, even of very restricted from, would be useful. A careful study of communication principles, of how people speak, hear, and comprehend, of how computers work, of the limitations of computer handling of natural languages or programming languages, and of the difficulties of proper electronic analysis and synthesis of spoken utterances suggests what might be rationally hoped for in immanent programs in speech communication with computers.

In this paper, we shall consider how studies in speech communication with computers will, on the one hand, extract from and depend upon, and, on the other hand, impinge upon and give impetus to, studies in related fields of: acoustics, phonetics, linguistics, language learning, speech and hearing physiology, psychology and perception theory, bionics, communication theory, human engineering, and computer science. We shall also consider the effects of such studies on age-old problems like aiding the physically handicapped and teaching second languages.

In section 2, the pros and cons of speech communication with computers are reviewed, particularly with respect to the specific advantages of speech as a man-computer communication modality in a spaceborne environment. Of all possible applications of speech recognition and synthesis devices, the use in spaceborne environments seems most encouraging, for reasons that will be discussed.

Recognizing the potential value of speech communication with aerospace computers, NASA has undertaken a comprehensive research program directed toward versatile man-computer communication by speech. This program is reviewed briefly in section 3.

What are the obtained and expected effects of this work in speech input/output (I/O) facilities for aerospace computers? The impact on future computer technology, on programs for computer handling of natural languages, and on the broad range of problems relevant to scientific understanding of human communication processes, is considerable, and is outlined in section 4. In section 5, some specific impacts on the work-a-day life of individuals and society are discussed. Mechanisms for aiding the physically handicapped and for improved methods of language training are expected to be valuable products of research on speech-handling devices.

2. Why Speech Communication with Computers?

It is surprising how little has been explicitly said to demonstrate concrete and practical reasons for speech communication with modern sophisticated computers. Few researchers have gone beyond the ostensive appeal and glamor of conversational machines to see or discuss the consequent practical and scientific advantages this new modality may offer.

The objective in this section is to survey some specific factors relevant to the question of why one should (or should not) have speech facilities in man-computer systems. Particular emphasis will be given to applications for manned space missions and other aerospace applications.

2.1 Natural Language Communication

The set of problems to which digital computers are being applied is rapidly expanding from the initial simple arithmetic operations to the more exotic and human-like tasks of abstract symbol manipulation, information retrieval, pattern recognition, language translation, theorem proving, game playing, etc.[15] In particular, this expansion of applications is expected to occur for spaceborne computers used on future manned space missions such as manned orbiting laboratories, extended lunar missions, and missions to Mars and other planets. Such spaceborne computers will not be just guidance and control calculating machines. They will undoubtedly be large elaborate multipurpose systems[8] concerned with performing elaborate guidance and control functions, standard mathematical manipulations, novel symbol manipulations, scientific experimentation functions, and more exotic "artificial intelligence" tasks of various forms.[15, 26, 27]

With the trend toward computers handling more intelligent and natural-to-human problems have come strong interest in and demand for natural language communication with machines. Some arguments for such natural language man-computer communication have been given elsewhere. (See, for example, (27) or (12).)

Speech is perhaps man's most natural, universal, and familiar form of communication. Speech communication is pleasant to the human, is learned at an early age, and is subsequently used and understood more universally than any other modality such as writing or typing.

Consequently, it is reasonable to seriously consider developing facilities for natural-language speech communication with computers, and, as will soon become evident, particularly with spaceborne computers. The extensive training and experience which the astronauts or other computer users have with spoken natural languages would thus be well utilized, special training in the use of other computer input/output (I/O) devices might be avoided, and wider and more effective use of spaceborne computers will be possible for all astronauts, regardless of technical background.

2.2 Aiding the Busy Computer User

Most of us are at least remotely familiar with the fantastic work loads which air traffic controllers experience. Under severe pressures and demanding circumstances they must absorb, process, and put out several varieties of information. Computers are being incorporated to ease the load and aid the controller to some extent. The importance and value of speech communication in their situations is apparent, and one might hope and expect that a computer with speech communication facilities could contribute markedly to the overall man-machine performance.

Similarly, there will undoubtedly be times during space missions when speech will be the most convenient or perhaps even the only way for communicating between the man and the computer. For example, suppose the astronaut should be piloting the ship or performing a guidance task where he, say, is "looking out the window" or monitoring visual displays. Both his hands and eyes are busy, and his ability to observe other (visual) computer outputs is restricted. Likewise, his ability to input to the computer by typing, handwriting on a graphical RAND tablet[11] or using other tactile devices would be restricted. Voice communication would then be useful as a substitute or augmentation for other I/O equipment.

Likewise, a variety of scientific experiments and tasks are to be expected in which the astronaut-scientist has his hands and eyes busy. He still may need to communicate with the computer. Speech I/O offers that option and added capability.

2.3 Communication Capacities and Multimodal Communication

This added modality, speech, thus allows communication where or when it might otherwise be impossible. It also provides increased channel capacity for communication with the computer during critical times when the astronaut and computer may be very busy.

The increased capacity for communicating information to and from the computer is quite significant. Obviously, the more I/O channels simultaneously available, the more information one can transfer in a given time span. Likewise, if there is a choice to be made between speech and some other channel, speech is likely to often come out ahead. Speaking surpasses writing or typewriting (or other of man's output modalities) with respect to speed and ease of information transfer. For example, Lea[27] has shown that to equal the rate at which speech conveys information, one would have to typewrite at over 100 words per minute. This speed is far beyond the ability of most typists.

Voice communication thus appears "better" from a channel capacity standpoint than any one other modality of man-computer communication. It also permits multimodal communication with the machine. Multimodal communication offers a versatility not readily possible with any one modality alone. Auditory information may be used in a supplementary fashion to augment information communicated by other modalities, such as in localizing specific parts of a graphical presentation. Multimodal communication also permits using one modality, such as speech, as a back-up channel in case some other I/O device fails.

Hence, multimodal communication with computers including speech I/O offers a flexible, more powerful, and more reliable total system for man-computer interaction.

2.4 Astronaut Mobility

Undoubtedly one of the strongest advantages of speech communication with computers is the possibility of physical mobility and the absence of any requirements for physical contact or specific orientation with respect to the computer. Conventional computer I/O devices such as switches, lights, punched cards, paper tape, teletype, X-Y plotters, and even the more recent RAND tablets[11] and cathode-ray tube (CRT) displays require the user to have physical proximity and restricted orientations with respect to the computer.

In contrast, speech I/O permits drastically improved mobility and consequent system flexibility. The human user may walk around the vicinity of the computer while still communicating with the machine. This is a strong point in favor of speech communication with computers wherever severe constraints on user position and orientation are to be avoided. In particular, this mobility will undoubtedly be useful in space flight and will provide more flexible man-computer interactions.

Also, if an astronaut is outside the spaceship in space, simple inclusion of a microphone in his spacesuit and a means for electromagnetic transmission back to the ship will allow his continued speech communication with the computer on board[24]. This is in marked contrast to the complications and inconvenience of using other I/O equipment under such conditions.

2.5 The Closed Spaceborne Environment

Since there is no matter to sustain the vibrations of sound in space, it is impossible for outside auditory noises to interfere with the man-computer speech link. Such advantages (while they may be overshadowed by interference due to noise sources aboard the spaceship) are in marked contrast to problems in research laboratories, factories, or wherever else speech I/O might be interferred with by uncontrollable outside noises.

One other advantage of the closed environment aboard a spaceship is that, for many foreseeable space missions, the number of people who will be using the voice link to the computer is quite small. One of the most difficult problems in speech communication with computers is automatic machine recognition of the same utterance spoken by different speakers. Each speaker has his own distinctive or ideosyncratic way of uttering any word, phrase, or sentence.

With only five or ten astronauts aboard a mission such as a Mars mission, the population invariance problem is reduced to that of handling this small set of voices. The computer could even to "tuned" or trained to each individual voice. The process of automatic speech recognition is thus considerably simplified by confining the speaking population to a small group such as the astronauts aboard a spaceship.

2.6 Some Problems With Speech I/O

Despite its many advantages, speech I/O has its drawbacks, too. These have been discussed in some detail in[27]. In brief they are as follows. Speech leaves no permanent record unless redundant storage devices are used. Some effort must be made to eliminate auditory noises which may interfere with the man-computer speech link. Spoken communication between man and computer will involve a complicated compromise between the human needs for high expressive power and naturalness in the language used, and the computer's needs for nonambiguous, mechanical languages of restricted form[20]. Complicated speech recognition and synthesis programs, as well as sizeable syntactic and semantic analyzers, must be stored and processed within the computer system. Finally, and perhaps the most immediately demanding of all, speech I/O peripheral devices must be developed to achieve versatile automatic speech recognition and smooth-flowing, natural speech output from the computer.

Techniques for resolving some of these problems are forthcoming from programs like those discussed in section 3.

2.7 A Summary of Factors Related to Evaluating Voice I/O

Figure 1 summarizes the many factors relating to the value of voice I/O. Human effects include the pleasure and naturalness of talking, the best use of the man's communicating abilities and extensive training, the need to communicate even with one's hands and eyes busy, and the desirable physical mobility permitted by speech. The comparison of the several channels for man-computer interaction favors use of the high-capacity speech channel and multimodal communication. For each of the channels, the languages used must represent compromises between human needs for high expressive power and naturalness versus computer needs for nonambiguous, restricted languages.

To process any reasonably versatile speech input or output, the computer system must incorporate peripheral devices for automatic speech recognition and synthesis. Also required are sizeable software programs for speech sound or word categorization, plus stored programs for syntactic and semantic processing of sentences or commands.

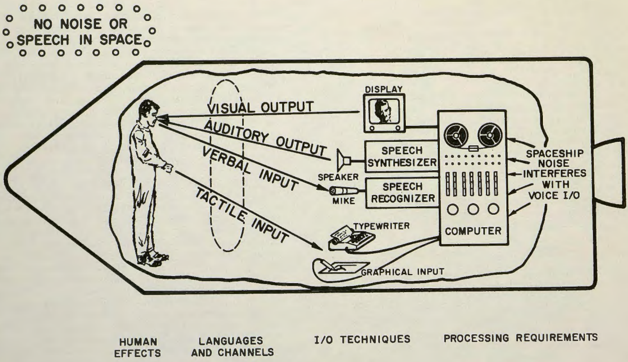


Figure 1. Factors Relating to Evaluation of Voice I/O

The development of I/O techniques or devices for automatic recognition and synthesis of speech is one of the most pressing problems in computer processing of speech. On the other hand, the processing requirements, such as size of computer storage and processing capabilities, will be directly relevant to how much bulk, weight, and complexity will be required in computers with facilities for man-computer voice communication.

3. What Does Speech I/O Require?

It is beyond the scope of this paper to discuss the many detailed problems involved in implementing speech communication with computers. The reader interested in detailed studies and surveys of work in automatic speech processing is referred to Flanagan's book[16] or the University of Michigan two-volume treatise on Automatic Speech Recognition[2], or Lindgren's survey[28] and references therein. A survey here of the general types of problems will, however, help show the scope and interdisciplinary efforts required for any measure of success. Study of the general problem of versatile man-machine communication, including speech, has led to an extensive program in speech I/O at NASA's Electronics Research Center.

3.1 Duplicating Human Communication

As illustrated in Figure 2, the basic problem of speech communication with computers is to functionally duplicate the behavior of the man-man communication link. Human communication is concerned with conveying mental or emotional concepts ("meanings" or "messages"[42]) from one person to others by means of agreed-upon symbols[13]. When a computer is to act as a speech output device, it in effect is duplicating the intelligent human speaker's process of converting from internal symbolic representations or concepts into articulated signals of continuous acoustic form. When accepting inputs spoken by the man, it is involved in duplicating the overall behavior of the human ear-brain sensory reception and perception system. This is illustrated in comparing Figure 2a with Figure 2b.

3.2 Requirements of Speech I/O Facilities

The important communication processes must be precisely understood if one is to mechanize them. The ultimate purpose of communication

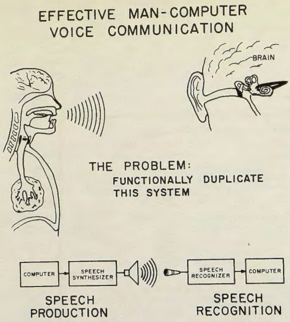


Figure 2. The Man-Computer Speech Communication Problem

(which may be simply stated to be to affect the expectations and future response patterns of the audience) must especially be understood and used to develop criteria for evaluating the effectiveness of a communication. The real overall criterion for judging the effectiveness of a communication is not simply to insure that transmitted signals are accurately received[42], or even properly categorized and assigned 'meanings'. Rather, it is that the communication yield the desired response from the receiver or audience.

Thus, in particular, the ultimate goal for computer input is to have the computer yield a desired response. That response might be a simple pre-determined reply by graphical display, teletypewriter, punched cards, spoken output, or what-have-you. Or, the response might be to merely store within its memory certain information, for future use; or it might be to locate and provide previously stored information. On the other hand, the desired response might be for the computer to make elaborate decisions based on the received information, or to compute some arithmetic function, such as a Fourier transform or matrix inversion. The possible examples of responses are as extensive as the set of problems to which computers might be applied.

Similarly, the goals for computer output might be to inform the human user, to request further information, to get the user to make decisions, correct previous errors, etc.

Recognition of these broad goals of communication leads to a realization of how much more complex and interdisciplinary is the problem of automatic speech processing than is generally recognized. Considering first the more difficult problem of automatic speech recognition, a typical comment is that by the late leading researcher Gordon Peterson:

> "Automatic speech recognition is primarily concerned with the conversion of the continuous functions of speech to a discrete symbolic representation." 37

It is certainly true that automatic speech recognition entails important continuous-to-discrete conversion processes. Speech is not simply spoken writing; rather, it is a continuous stream of sounds or mechanical vibrations, not easily segmented into words or other discrete symbols. The acoustical wave produced by the speaker and received by the speech recognizer is a continually-varying pressure function which can take on a continuum of pressure values. Yet, not every possible wave shape is linguistically distinguishable from another. Very slight changes in the acoustical wave will not generally change the meaning of an utterance. Such slight changes (due to different configurations of the individual speaker's vocal tract, etc.) do not signal changes in desired responses. There is thus a certain quantization process which goes on in the recognition of speech, so that "similar" sounds are grouped together into one sound class. This categorization is the subject of the phonological (allophonic, phonetic, phonemic, and prosodemic) component of the linguistic description of a language.[23]

Although the process of transforming continuous speech signals into discrete codes is the recognized core of speech recognition, it is not the whole story. The meanings, or messages, intended to be conveyed, and meanings actually aroused in the listener, are as relevant to man-computer communication as they are to man-man communication[31]. Some consideration must be given to the semantics of a communication.

Speech recognition cannot be readily accomplished without adequate assignment (from the sound structure, or phonemic, level to the sentence level) of a grammatical structure and meaning. The grammar, or syntax, which tells how symbols or words or such may be combined together to yield acceptable ("grammatical") utterances, must be incorporated into the computer programs for the generation or recognition of utterances.

In view of these comments, and the value of considering the ultimate purpose of communication expressed in terms of the altered internal state and responses of the computer, we may describe a general model of automatic speech recognition which relates the continuous-to-discrete conversion process with linguistic analysis, semantic assignments, and computer responses. Speech recognition, whether by man or by machine, is the process of transforming the continuous acoustic speech signal into discrete symbolic representations which may be assigned meanings and which, when comprehended, may be used to affect responsive behavior.

3.3 A General Speech I/O Model

One general model of the processes of automatic speech recognition (speech input) and synthesis (speech output) is illustrated in Figure 3. According to this model, the process of speech recognition begins with a means for extracting, from the highly-redundant continuous speech signal, some important "information-carrying" speech parameters[37]. One frequent technique is to extract the 'formants', or spectral concentrations of energy in the frequency spectrum of the speech signal and to track their time changes through phonemes, syllables, and longer utterances[7, 14, 44]. Other techniques have been developed for extracting just those features or parameters that are necessary to establishing what speech sounds were said[17, 27, 28, 33, 43].

The speech parameter extraction problem is perhaps one of the most demanding problems in automatic speech recognition. It concerns the fundamental question of what basic measurements must be made on the speech signal to permit establishing what was said. According to Lea[27, p. 194] there are three general viewpoints which can guide our selection of meaningful measurements. The speech production viewpoint suggests important "information-carrying" parameters based on the way the speech was produced by the human speaker. The speech perception viewpoint seeks to establish relevant parameters in light of the characteristic manner in which men perceive speech. The third viewpoint suggests that we duplicate the human sensory reception processes to yield a set of parameters equivalent to those extracted by the auditory sensory system (the sensory reception viewpoint). To these three viewpoints, we may add the more naive acoustic or physical viewpoint, which says that parameters will be selected on the basis of general signal detection techniques and physical processes, ignoring the nature of the source and destination of the speech signal.

A system for experimentally comparing various speech parameter extraction techniques, based on these viewpoints, was discussed by Lea[27] and is being implemented at NASA. The importance of this work to this paper is, however, simply to illustrate the many fields that must be tapped to facilitate speech recognition system design. Understanding of how people speak (including how concepts are converted to neural control signals for speech articulator muscles, how speech is articulated, how articulations relate to resulting acoustic parameters like spectra and formants, and how articulatory correlates may be determined from the acoustic signal) is important to parameter extraction and to subsequent determination of sound categories.

Similarly, understanding of how the human ear-brain system achieves speech recognition (including how the acoustic signal is physiologically transformed by the outer ear, oscicles, and inner ear, how the acoustic signal is encoded into cochlear nerve outputs, and what subsequent neural processing is done on these sensory inputs to the brain) may give us helpful hints about relevant parameters from the sensory reception viewpoint. Likewise, understanding of human perception of speech (including psychometric studies, listener tests, etc.) may provide perceptual cues about what is really important to the categorization of speech signals.

Once significant parameters have been extracted from the speech signal, the parameter patterns must be recognized or classified. There are many acceptable techniques for pattern classification when the number of distinguishable patterns is not large[19, 28, 35] and they will not be discussed here. Rather, we may merely observe that the recognition techniques will, in the most general case, consist of an iterative process as illustrated in Figure 3. Recogniton of some speech segment, such as the phoneme, syllable, or word, will be achieved based on the parameter pattern in the segment. (Considering the difficulty previously experienced in phoneme recognition, in particular, this sound category recognition may be incomplete, such as only indicating whether the sound was a vowel, stop constant, diphthong, or such general segment type rather than exactly which of the 40 or so English phonemes was said, etc.). Sequences of these recognized sound categories might then be recognized as composing larger units such as syllables, words, phrases, or sentences. This sequence recogniton may thus involve consideration of coarticulation in the language, what sounds can follow what, how preceding and following sound contexts may change particular sounds, what morphophonemic constraints there are on sequences in the language, what the syntax or grammar says about possible sound sequences, sentence structures, etc. Finally, semantic recognition based on the recognized sound structure and grammatical structure will determine what meaning was conveyed and what replies or machine responses are to be invoked.

Results out of any recognizer block may be fed back (see Figure 3) to previous blocks to adjust future categorizations based on previous decisions. Thus, even though preliminary predictions of exact phoneme categories may not be possible on a first pass, use of fedback information about decision context may facilitate final decisions on phonemes after several iterations.

Figure 3 illustrates the central role played by the linguistic description in automatic speech recognition. The linguistic description specifies what parameters are really important to determining what linguistic utterance was spoken, what sound categories are relevant to a particular language, what sequences may occur in properly-formed sentences of the languages, and what meanings are to be assigned to utterances.

Figure 3 also illustrates corresponding processes in speech synthesis. What response is to be obtained determines what message must be conveyed, and that message or meaning must be encoded as a grammatical utterance, and, in turn, as a sound sequence which controls a speech sound synthesizer.

3.4 A Research Program for Speech Communication With Computers

Preliminary research into the various processes involved in speech communication with computers, and the corresponding human communication processes, has led this author to a comprehensive program plan for building up versatile capabilities in man-computer speech at NASA Electronics Research Center. In order of precedence, the five major areas of research planned are concerned with: (1) goals, requirements, and applications; (2) automatic speech recognition; (3) continuous speech processes; (4) speech output; and (5) linguistic models.

3.4.1 Goals. The discussion of the value of speech communication with computers, as presented in reference 27 and reviewed in section 2 above, is a result of work on the goals, requirements, and applications of speech I/O. Another example requirements program (which has just recently been undertaken under NASA sponsorship) deals with the acceptability to the human user of restricted forms of speech I/O. Fully versatile, unrestricted speech I/O is not immediately forth-

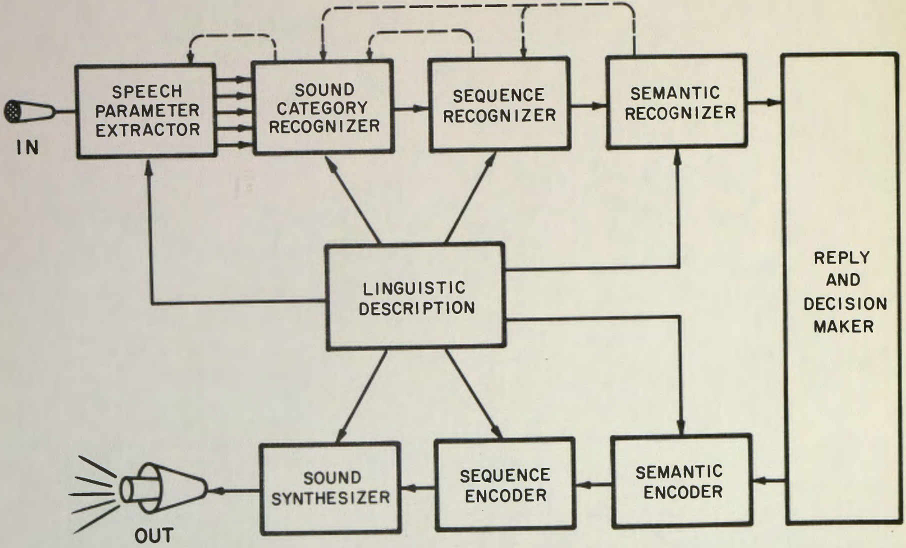


Figure 3. A General Model of Speech Processing

coming[7], and one might well ask how restricted a system can be and still be useful. What constraints on the spoken vocabulary, syntax, semantics, verbal system response times, etc., do not severely impair the success of the man-machine system? The answer to this question will bear directly on what features a minimal speech I/O device must include.

3.4.2 Automatic Speech Recognition. Research in automatic speech recognition is the more difficult problem requiring longer lead time before successful implementation. Speech parameter extraction, more fundamental and more difficult than development of recognition algorithms, needs particularly careful attention. Besides conventional work in speech spectra extraction[7], NASA is sponsoring research at the University of Illinois on fundamental speech properties[3], and is implementing the "SPEECS System" (reference 27, pp. 193-6) for experimentally comparing parameter extraction techniques. Work on realistic recognition systems has included a successful system for recognition of words from vocabularies of 50 to 150 words[7] and a pilot study of optimal damped simusoidal analysis procedures in speech recognition[9].

3.4.3 Continuous Speech Processing. Previous efforts in automatic speech recognition and synthesis by machine have been confined primarily to the handling of isolated spoken words or phrases. Recognition techniques used have viewed the word or phrase as a conglomerate block or pattern. Such approaches are not directly applicable to any practical recognition of continuously-flowing human speech or versatile, structured sentences and non-discrete utterances. The proper recognition or synthesis of continuous speech demands a comprehensive technique for relating the continuous speech signal to its discrete representation as a string of phonetic segments, syllables, words, phrases, breath groups, or other such concatenations.

This requires careful study of continuous speech processes. Problems of prime importance in continuous speech recognition include how an utterance may be segmented into recognizable portions (the segmentation problem), how to make the recognition process insensitive to segmentation errors, and how to improve prosodic aspects, which are so important to the flow of speech.

The segmentation of the continuous signal into useful segments, and the application of speech prosodies to such segmentation and to the synthesis of natural-sounding, smooth-flowing speech, strongly demonstrate the need for basic studies of the prosodies (or suprasegmental aspects) of speech if one is to be able to systematically attack the problems of continuous speech input to, and output from, a computer. NASA has recently begun a sponsored research program at the Speech Communications Research Laboratory to study the prosodies of speech. Work on segmentation techniques will follow.

3.4.4 Speech Synthesis. As argued above[27], earliest research should be directed more toward advancing the technology of computer recognition of speech than to the easier problem of computer synthesis of speech. However, after the proper lead time is given to recognition and continuous speech processes, the problem of synthesis (i.e., of producing desired speech outputs from discrete symbolic representations in the computer) must be considered. One naive approach to speech output is the storage and retrieval of pre-recorded messages. This technique, using digital encoding of inputted speech and storage on the magnetic disk of a time-shared computer, has been implemented in a simple but novel form at NASA/ERC. A more versatile form of speech output would be automatic speech synthesis, whereby complex continuous-flowing utterances were formed out of small sound segments.[32, 41]

3.4.5 Linguistic Models. A major long-range problem area concerns linguistic models. It is the linguistic description of a language (whether it be a natural language like English or a computer programming language) which establishes when speech sounds are to be considered the same or different. The linguistic description also relates the sound structure of phonetic and prosodic aspects of speech to the grammatical or syntactical structure, the meaning of an utterance, and the purposive use of utterances in communication. Thus, adequate linguistic descriptions, incorporating phonetic, phonemic, morphophonemic, syntactic and semantic models, must be understood and specified in precise form for computer recognition and production of speech.

The overall program in speech communication with computers, as described here, involves more than the explicit characterization of goals, automatic speech recognition techniques, continuous speech processing, speech output techniques, and precise linguistic models. Considerable basic research into human communication processes like speech articulation, acoustic phonetics, sensory reception and hearing processes, perception and psychology of hearing, phonemics, syntax, semantics, etc., are intrinsic to achieving such practical goals.

4. The Scientific Impact

In the introductory sections of this paper, the desirability and many advantages, and, indeed, it might be argued, the "necessity" of speech communication with computers was discussed. Those who have accepted the trend toward more versatile computer facilities and the specific advantages of speech I/O might then readily accept the necessity of a program such as has been presented in section 3. A comprehensive program of the inter-disciplinary form demanded for speech communication with computers will tap the resources of many fields related to human communication, speech and hearing, and computer sciences. We might also expect such study to yield contributions back to the fields it taps. This is indeed the case.

If necessity is the mother of invention, it is also a prolific progenitor. The necessity of attaining a particular goal (such as speech communication with computers) most generally produces not only the inventive solution to the initial need, but also yields the discovery of many other new "necessities" and the solutions to many related problems of practical and theoretical form. Thus, e.g., research and development work on specific problems of speech I/O has a decided impact on the science and technolgy of the day.

In the present context, we can only sketch the variety of effects which speech I/O research has had and will have on various disciplines. Detailed discussions and lists of examples are beyond the scope of this paper.

The most straight-forward and immediate impact of research on speech communication with computers is in advancing the potential of achieving the original goal of versatile man-computer interaction. As argued in section 2, voice communication with computers can be a very valuable addition to the man-machine interface, allowing new dimensions of user mobility, increased communication capacities, natural-language communication, the flexibility and reliability of multimodal communication, and proper use of the user's training and experience with speech and spoken language. These and other advantages are particularly manifested in aerospace applications, such as manned space missions[27] and astronaut extra-vehicular maneuvering units.[24]

But speech I/O is also quite relevant to other man-machine contexts. With the recent congestion of airways and airport corridors, and subsequent delays and threats to safety, the problems of air traffic control are particularly acute. Since speech communication has always been a dominent part of the controller's activity, and recent trends are toward the increased use of computer aids for the busy controller, it is reasonable to expect the speech communication with computers might be appropriate for air traffic control facilities. Here speech I/O research would have a direct impact on a major technological crisis of our day. Similarly, speech I/O facilities may be valuable tools for post office zip code readers, airline reservation facilities, banks, synthesized and recorded telephone replies, etc. Indeed, support of such applications has been provided.

One interesting application of speech I/O which is relevant to many present-day computer installations is as an adjunct to graphical display systems. Display systems often have actual small keyboards of buttons, or a small set of virtual "buttons" on the CRT display, to control modes of display, character recognition, and other gross characteristics of display. Rather than selecting and pushing the appropriate button, or activating the appropriate "virtual button" of the display with a light pen, one can speak a vocabulary word to the computer to control its mode of display. This multimodal man-machine interaction is particularly satisfying and effective.

An obvious by-product of sponsored research and development on speech communication with computers is the development of a substantial technology base and groups of research teams necessary to solve any related problems that may arise. On short notice, such a technological workforce is available for application to immediate needs in speech and man-machine related problems.

The precise formulations and techniques required for automatic processing of speech also serve a strong and far-reaching corrective purpose in speech and communications research. For example, linguists have for years presented linguistic models for the articulatory, phonetic, prosodic, phonemic, morphophonemic, and syntactic structures of natural languages. Such models have generally been informal statements of rules for sentence construction, acceptable sound structures, what distinguishes phoneme from phoneme, etc. The informal models depend heavily upon intelligent interpretation and application by humans. They are not suitable for direct implementation on machines. Consequently, with the recent extensive attempts[2, 15, 20, 30] to use computers for a variety of natural-language processes, such as mechanical translation, information retrieval, solving problems requiring "artificial intelligence", and "understanding" of spoken inputs and consequent production of verbal replies, there has arisen a need to develop a mechanistic, "non-creative" theory or model of natural language. Automatic speech processing requires defining exactly and algorithmically what is meant by such conventional linguistic constructs as "phoneme categories", "phonetic juncture", "meanings", etc. When such has been attempted, the repeated experience has been that previously acclaimed linguistic models have proven inadequate and in need of fundamental revision and improvement.[2, 23, 28] The search for acoustical and articulatory correlates for linguistic constructs has been particularly revealing in this regard. Mechanical tracking of phonation or voicing, other 'distinctive features', prosodemic elements, and other phonological properties has disclosed specific errors and inadequacies of traditional linguistic descriptions, and in some cases revealed ways in which such models may be improved.

Similarly, the search for the fundamental speech properties most relevant to automatic speech recognition and synthesis has demonstrated some inadequacies of previous attempts at speech analysis by spectral analysis. Such research has also suggested some alternative explanations for the sensory reception operations in the human ear.[16, 43]

The demands of automatic processing of speech have, in general, given impetus to expanding research on the basic aspects of human communication, speech, and hearing. Researchers working on automatic speech processing have repeatedly emphasized the need, and offerred new techniques, for futher studies in speech physiology, acoustic and articulatory phonetics, phonemics and pattern classifications, "information-carrying" parameters of speech, perception of speech and other signals, speech prosodies and what makes for the "naturalness" of a language, information theory and other aspects of human communication.[14, 19, 28, 29, 32]

Lest anyone should fail to realize the general scientific and sociological impact of such studies he should reflect carefully on the universal significance of speech and related subjects to all communicating humans. Language and communication have played a central role in the advance of science, the rise and fall of societies and institutions, and the well-being of mankind. Either directly or indirectly, speech, language and communication are subjects of concern to almost every type of scholar or scientist. Besides the obvious use of such in the presentation of any scientific theory or experimental results, speech, language, and communication are of specific interest to any scholar whose work depends directly upon human actions. Linguists, of course, owe their very profession and problems to the existence of human communication and common 'languages" of agreed-upon symbols. Communication theorists, on the other hand, find studies of human communication to be an important area of possible extension of their theories of technical signal transfer. Psychologists may find interest in verbal behavior as an important "stimulus-response" process, and many interesting questions might be raised about how a person perceives, why a person "decides" to engage in the generation of utterances, how received messages relate to subsequent response behavior (verbal or otherwise), how utterances evoke emotional responses, etc. Philosophers may be intrigued by the logic involved in communication, and, indeed, it has been said[1] that an essential problem of philosophy is the meaning of words and expressions used in arguments. From another viewpoint, physiologists and cyberneticists can find interest in the description of neurophysiological processes going on in the brain during speech and hearing. Also of interest are the observable physical processes of coordinated muscle movement, acoustic vibration of the conducting medium, and sensory reception of the traveling auditory signals.

Such a listing could continue, demonstrating widespread and diversified interest in communication, languages, and speech. Similar arguments could be made for the growing importance of computers which are readily accessible to untrained users. It is in the light of such widespread importance that the contributions of work in speech communication with computers (and, specifically, work in aerospace applications of such) must be evaluated.

5. The Social Impact

The previous remarks about the impact of speech I/O work on the understanding of human communication illustrate the widespread social impact that may be expected. Yet, the man on the street, and even many workers in various related technological disciplines, may not be aware of any overt changes or improvements which one can readily trace back to the specific efforts in speech communication with aerospace computers. Here we shall briefly consider a few important applications which, while they certainly do not owe their successful outcomes solely to studies in speech I/O, do illustrate how such studies are one strong impetus and aid to new developments.

Some studies obviously may be expected to benefit from, and add their own contributions to, studies in speech I/O. Among such studies are ones concerned with: reconstruction of deep-sea divers' "helium speech"; automatic speech recognition or authentication; rate-controlled speech; speech privacy and encrypted speech, using vocoder-like signals; audio response units; and various sensory aids to help the deaf, blind, and those otherwise physically handicapped.[29]

5.1 Sensory Aids

One of the applications of speech studies which has perhaps the strongest appeal and social visibility is the development of sensory aids for the handicapped. Past speech research has yielded concepts and devices relevant to: the use of accelerated speech to aid the blind; the use of slowed-down speech to perhaps help the mentally retarded; the possibility of tactile vocoders for the deaf; frequency translation and compression to help the partially deaf; artificial larynxes for those who have lost their ability to produce voiced sounds in the larynx[39; p. 170]; etc.[29]

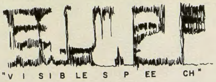
Past aids of such forms make one hopeful that present and future speech research will likewise result in many worthwhile applications and devices. However, before considering specific applications to the aiding of physically handicapped, it is worthwhile to recall the precaution taken by earlier speech researchers in applying their research to the problems of the physically handicapped.[40; p.22] In his zeal to sell or publicize his ideas to the general public and to the blind, deaf, dumb, and others who would be inclined to welcome new aids with open arms and overzealous hopes, the speech researcher would do well to remember the discouragement and damage that can incur from devices which fail to achieve all he promises or hopes. We may all welcome the advent of new successful aids like the hearing aid or artificial larynx, but caution is in order when presenting or considering experimental devices and techniques which have not been developed to the stage of proven performance.

Besides some substantial contributions to the realization of the aids discussed previously, we may note a few other new developments which are related to speech I/O research and are worthy of careful, judicious study by those interested in aids to the handicapped. There is, for example, the possibility of using large computer systems and versatile speech synthesis facilities to provide aural book reading for the blind and others unable to read. Similarly, the blind and others unable to communicate with a computer through standard input devices would find speech I/O useful as a means for entering computer programs, or perhaps (if the day should ever come when such is feasible) as a phonemic or voice-operated "typewriter".[36]

Perhaps even more promising is the potential application of graphical displays of speech. The acoustical speech signal is converted into a visual pattern which might be displayed on a plotter, strip-chart, cathode-ray-tube display, or such. Such "visible speech" has, in several different implementations, proven of considerable value in a variety of speech studies, including as aids to the deaf.[25,29,39,40] Since the development of the spectrograph in 1947, considerable work has been done on its use in training the deaf to "read" visible speech patterns corresponding to spoken utterances.[40] A sketch of the form of display the spectrogram achieves is illustrated in Figure 4(a). The person reading the pattern must follow the relative positions of the dark bars through the word, and read the corresponding speech sounds the patterns represent.

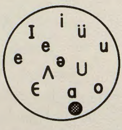
Another promising visual display of speech, which has resulted from work on computer-controlled graphical displays and research on important speech parameters and the articulatory correlates of speech signals, is a method for portraying articulator motions in a picture (mid-sagital view) of the human vocal tract.[5, 21, 29] A typical picture of the display at one specific time might be as in Figure 4(b), where the speaker is shown to be in the process of articulating an "ah" sound, as in'father'. Such displays of "visible articulation" might be considered as an extension of traditional "lip reading" by the deaf.

Other displays of speech have involved such techniques as displaying the values of resonant frequencies (formants) of the vocal tract as the coordinates of an X - Y display (F_1 -- F_2 plots[10]) or variously displaying the pitch or other selected parameters [4, 5, 6, 38] of the speech signals. These techniques have been directly based on the results of research on automatic speech processing. Another potentially meaningful type of display might be any of several ways of displaying the yes-no properties of "distinctive features".[19, 38]



(a) SPECTROGRAPHIC DISPLAY OF THE SPOKEN PHRASE "VISIBLE SPEECH"

(b) ARTICULATORY DISPLAY OF SPEECH, SHOWING CONFIGURATION FOR "A" AS IN "FATHER".



(c) VOWEL-CIRCLE DISPLAY, SHOWING "DOT" (CROSS HATCHED) FOR SPOKEN "A" AS IN "FATHER".

Figure 4. Visual Displays of Speech

Other work has considered the use of color displays of relevant speech parameters, or color instead of shades of darkness in spectrograms.[40] Such concepts are more feasible now that recent research has been directed toward useful computer-controlled color displays.[6, 45] Work with one such color system has included studies of deaf children reading the color display of speech.[6]

Figure 4(c) illustrates a third type of visual display of speech, which has resulted from NASA-sponsored studies in speech perception and machine recognition of speech.[46, 47] By a complex analysis of incoming speech, the system yields a display as in Figure 4(c). When a specific vowel is spoken, the trace ("dot") on the cathode ray tube goes to the position of the particular vowel being spoken. The symbols on the scope face indicate the phonetic category of the spoken vowels.[46] Consonants and transitional sounds are shown as transitions from place to place on the display. Preliminary tests have indicated that deaf and mentally retarded persons can learn to improve their own speech articulation, as well as read other people's speech, using such displays.[34, 47]

Such visual displays of speech are one valuable attempt to resolve a major problem concerning the deaf. That problem was described by Gunnar Fant[29;p.84] as "how we can develop speech feedback mechanisms internally via kinesthetic and proprioceptive mechanisms and externally via auditory, visual and tactile recording."

5.2 Language Learning

The possibility of providing feedback to help the deaf to speak better, despite their inability to hear and accordingly correct their own speech abnormalities or errors, illustrates the training possibilities of visual displays of speech. However, there is another major application of visual displays to the speech training of individuals. This is the use of visual displays in teaching students to properly pronounce utterances of a language, particularly a second language they are just learning to speak. Given a well-designed visual display of speech, an instructor (or a machine itself) may show the distinction between correct pronunciation patterns and the student's pattern, and train the student to make the appropriate corrections.

Any of the display techniques discussed previously for providing aids to the physically handicapped are also potential candidates for aids to language learning. For example, one form of display which has been considered[5, 21] is to show, on a computer-controlled graphical display, a mid-sagital view of an ideal vocal tract configuration for the sound being spoken and a superimposed view of the speaker's (student's) vocal tract, as derived from the received acoustic signal. The vowel circle of Figure 4(c), with "correct" and "student's" pronunciations both displayed, is another alternative, and there are many more. The problem in selecting and building such displays is how best to display the importantparameters of speech so as to yield the most successful guide to what is wrong with the student's pronunciation and what must be done to correct it. Research on automatic processing of speech and what are the really important parameters of speech signals has offered, and will offer, valuable suggestions about appropriate displays.

The import of such aids to language learning is substantial. Most language teachers, and those familar with the conventional techniques and problems of language laboratories, will agree with Gloria Cooper, who said[4; p. 59]:

> "... the most potent aid to linguistic mastery is practice in producing utterances which can be corrected immediately."

The student's training in his first language makes it difficult for him to hear the distinctions appropriate to another language and the errors he makes. Determining an effective mode of information feedback for the student is a basic[3] aspect of the language training situation.

Immediate information feedback of correct and incorrect pronunciations is needed in order to correct the student's inappropriate responses before they become habits. Also valuable is some information about the nature of pronunciation errors and the direction in which the student must alter his pronunciation or articulation to achieve correct results. Such feedback is not readily available in conventional language laboratories, and is generally available only through tutors. A sophisticated, time-shared, on-line computer system and display may thus simulate a private tutor. This is particularly valuable during a time of shortage of qualified teachers.

Work on such computer-controlled aids to language learning is proceeding at various academic and research organizations. One study, at Bolt Beranek and Newman research facilities, has made substantial use of a system originally sponsored by NASA for computer recognition of speech.[4, 5]

Not only do computer-controlled visual displays of speech aid such language learning studies, but many other aspects of research on automatic speech processing are expected to provide techniques for improved speech training. One relevant aspect of the relationship between language learning and speech communication with computers is the fact that in both cases one strives for error-free sentence production. In both instances, then, a complete and accurate understanding of the phonology and syntax, and, indeed, all aspects of linguistic descriptions, is needed. Knowledge of what (prosodies, sound structure, etc.) makes speech natural and "speech-like", which is pertinent to automatic speech synthesis, may be useful in languagelearning. The selection of the most important speech parameters is common to both problems.

These contributions to the development of language training aids, and those to the aid of physically handicapped, are only a few of what appears to be a wide range of applications of the results of speech I/O research. When coupled with the scientific contributions to the understanding of human communication processes and the value and many advantages of speech communication with computers (particularly for aerospace applications), they all strongly encourage undertaking challenging programs directed toward providing versatile man-machine communication by speech.

References

1. Allston, W. P. The Philosophy of Language. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964.

2. Automatic Speech Recognition, Two-Volume Course in Speech Communication and Machines, University of Michigan Engineering Summer Conference, Report 6310, 1963.

3. Babcock, M. Speech Characterization; i.e., Analysis, Synthesis, and Recognition. In Accomplishment Summary, BCL Report No. 68.2, Biological Computer Laboratory, University of Illinois, Urbana, June, 1968, pp. 125-142.

4. BBN. Research in Second-Language Learning, Semiannual Technical Report No. 1, ARPA Order No. 890, BBN Report No. 1530 to ARPA, Air Force Office of Scientific Research, Arlington, Virginia, June, 1967.

5. BBN. Research in Second-Language Learning, Semiannual Technical Report No. 2, ARPA Order No. 890, BBN Report 1575 to ARPA, Air Force Office of Scientific Research, Arlington, Virginia, December, 1967.

6. Benninghof, W. J. A Functional Analogy Between Speech and Color Perception and Its Implementation for Sensory Replacement, PhD. Dissertation, Department of Electrical Engineering, Northeastern University, May, 1968.

7. Bobrow, D.G. and Klatt, D.H. A Limited Speech Recognition System, AFIPS PROC. of 1968 Fall Joint Computer Conference, Vol. 33. Washington:Thompson Book Co., 1968pp. 305-318.

8. Burke, T. and Wang, G.Y. Spaceborne Multiprocessing Organizations, Proc. 1966 WESCON, Session 9, pp. 1-6.

9. Colomb, R.M. and Roper, R. Optimal Orthonormal Bases for Spaces of Decaying Sinusoids; Final Report, Contract NAS12-684, prepared by Systems Research Laboratories, Dayton, Ohio, for NASA/ERC, Cambridge, Mass., December, 1968.

10. Davis, K. H., Biddulph, R., and Balashek, S. Automatic Recognition of Spoken Digits, J. Acoust. Soc. Am., Vol. 24, 637-642, 1952.

11. Davis, M. R. and Ellis, T.O. The RAND Tablet -- A Man-Machine Graphical Communication Device. AFIPS Conference Proc., 1964 FJCC, Vol. 26, Part I.

12. DEACON Project. DEACON Breadboard Processing, Report AD-612 165, TEMPO Division, General Electric Co. Santa Barbara, California, 1964. Available through Clearinghouse, U.S. Dept. of Commerce, Springfield, Va.

13. Dean, H.H. Effective Communication. New York: Prentice-Hall, Inc., 1953.

14. Fant, G. Acoustic Theory of Speech Production. The Hague: Mouton and Co., 1960.

15. Feigenbaum, E. A. and Feldman, J. Computers and Thought. New York: McGraw-Hill, 1963.

16. Flanagan, J.L. Speech Analysis, Synthesis, and Perception. New York: Academic Press, 1965.

17. Focht, L. R. The Single Equivalent Formant, Proc. IEEE Intern. Commun. Conf., Philadelphia, Pa., June, 1966.

18. Halle, M. Phonology in Generative Grammar, The Structure of Language (Fodor and Katz, Eds.). Englewood Cliffs, New Jersey: Prentice-Hall, Inc. , 1964, pp. 324-333.

19. Halle, M. and Stevens, K. N. Speech Recognition: A Model and a Program for Research. The Structure of Language (Fodor and Katz, Eds.). Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1964, pp. 604-612.

20. Hays, D. G. Readings in Automatic Language Processing. New York: American Elsevier Publ. Co., 1966.

21. Henke, W. Dynamic Articulatory Model of Speech Production Using Computer Simulation, PhD Dissertation, Massachusetts Institute of Technology, September, 1966.

22. Holmes, J. N., Mattingly, I.G., and Shearme, J.N. Speech Synthesis by Rule, Language and Speech, vol. 7, p. 127, July-Sept., 1964.

23. Katz, J. and Postal, P. An Integrated Theory of Linguistic Descriptions. Cambridge, Mass.: M.I.T. Press, 1964.

24. Kelley, T.P., Martin, J.T., and Barger, J.R. Voice Controller for Astronaut Maneuvering Unit. RCA Technical Report AFAL-TR-68-308, to Air Force Avionics Lab., Wright-Patterson Air Force Base, Ohio, October, 1968.

25. Kidder, C.W. Vocal Physiology and Bell's "Visible Speech". Boston: Publ. by author, Emerson College of Oratory, 1923.

26. Lea, W. A. Some Arguments Favoring Non-Conventional Types of Computers, NASA Technical Memorandum TM X-1544, March,1968.

27. Lea, W. A. Establishing the Value of Voice Communication with Computers, IEEE Trans. on Audio and Electroacoustics, Vol AU-16, No. 2, June, 1968, Pp. 184-197.

28. Lindgren, N. Machine Recognition of Human Language (A three-part survey), IEEE Spectrum, March - May, 1965.

29. Lindren, N. Directions for Speech Research (A report on a panel discussion of leading speech researchers), IEEE Spectrum, March, 1968, pp. 83-88.

30. Locke, W. N. and Booth, A. D. Machine Translation of Languages. Cambridge, Mass.: M.I.T. Press, 1955.

31. MacKay, D.M. The Place of "Meaning" in the Theory of Information, Information Theory, (E. C. Cherry, ed.), Buttersworths, pp. 215-225, 1956.

32. Mattingly, I.G. Synthesis by Rule of Prosodic Features, Language and Speech, vol. 9, p. 1, Jan.-Mar., 1966.

33. Mundie, J.R. and Moore, T. J. Use of Cochlea Patterns in Speech Analysis. Presented to 76th Meeting of Acoustical Society of America, Nov., 1968, paper L4.

34. NASA-Cambridge Project May Aid Deaf, NASA Electronics Research Center, Cambridge, Mass., News Release 67-39, Nov. 24, 1967.

35. Nillson, N. Learning Machines. New York: McGraw-Hill, 1965.

36. Olson, H.F. and Belar, H. Phonetic Typewriter, III, J. Acoust. Soc. Am., vol. 33, pp. 1610-1615, 1961.

37. Peterson, G. Automatic Speech Recognition Procedures. In Volume II, Automatic Speech Recognition. Univ. of Michigan Press, Ann Arbor, 1963, pp. Ha83-102.

38. Pickett, J.M. Recent Research on Speech-Analyzing Aids for the Deaf, Reprints of 1967 Conference on Speech Commun. And Proc., AFCRL, Office of Aerospace Research, 1967, p. 192.

39. Pierce, J. R. and David, E. E. Man's World of Sound. New York: Doubleday and Co., 1958.

40. Potter, R. K., Kopp, G. A., and Kopp, H. C. Visible Speech. New YOrk: Dover Publ., 1965. From the original D. VanNostrand edition of 1947.

41. Rabiner, L. R. Speech Synthesis by Rule: An Acoustic Domain Approach, Proc. 1967 Conf. on Speech Commun. and Processing. Bedford, Mass.: AFCRL, Office of Aerospace Research, 1967, paper A8, pp. 39-51.

42. Shannon, C. E. and Weaver, W. The Mathematical Theory of Communication. University of Illinois Press, Urbana, 1962.

43. Stewart, J. L. Frequency Filtering Processes in the Human Ear, Report by Santa Rita Technology, Menlo Park, California, Jan, 1966. Also refer to related articles by the same author in Biological Proptypes and Synthetic Systems. New York: Plenum Press, 1962.

44. Thomas, I. B. The Significance of the Second Formant in Speech Intelligibility, Technical Report No. 10, AF Grant 7-66. Urbana: Biological Computer Laboratory, University of Illinois, July, 1966.

45. Walter, C. M. Color -- A New Dimension in Man-Machine Graphics, Proc IFIPS Congress, 1965.

46. Yilmaz, H. A Program of Research Directed Toward the Efficient and Accurate Machine Recognition of Human Speech, Final Report No. 1, Contract NAS12-129, A.D. Little Report to NASA/ERC, Cambridge, MA., Dec,1966.

47. Yilmaz, H. A Program of Research Directed Toward the Efficient and Accurate Machine Recognition of Speech: A Theory of Speech Perception, Final Report No. 2, Contract NAS12-129, A. D. Little Report to NASA/ERC, Cambridge, Mass., Nov., 1967.