NOISE-CON 97
The Pennsylvania State University
University Park, Pennsylvania
1997 June 15-17

RELATIONSHIPS BETWEEN SPEECH INTELLIGIBILITY AND OBJECTIVE ACOUSTICAL PARAMETERS OR ARCHITECTURAL FEATURES IN CATHOLIC CHURCHES

António P. O. Carvalho, Ph.D. (1), António Morgado (1) and Luís Henrique (2)

(1) Acoustical Lab., Dep. Civil Eng., College of Eng., U. Porto, R. Bragas, 4099 Porto Codex, Portugal (fax: 351.2.2041940, email: carvalho@fe.up.pt) (2) School of Music and Performing Arts, Polyt. Inst. Porto, R. Alegria, 4000 Porto, Portugal

INTRODUCTION

This study is part of a research program initiated in 1991 by the author at the University of Porto (Portugal) and University of Florida (USA). The aim of the project is to explore methods to evaluate, predict and preview the acoustical qualities of churches. The program has included two major components to date:

- *Objective studies of existing churches* - Measurements were taken in 41 Portuguese Catholic churches, at multiple locations in each room. Several objective acoustical parameters were measured (RT, EDT, C80, D50, TS, L, BR_RT, BR_L, and RASTI) (Carvalho 1994).
- *Subjective studies of existing churches* - This has included both evaluating live musical performances in 36 churches and speech intelligibility testing. This work is characterized by the use of a sample of listeners, evaluation of several locations in each room, assessment of many rooms and comprehensive statistical analysis of the data (Carvalho *et al.* 1996).

This paper presents a report concerning relationships between speech intelligibility and objective acoustical parameters and with the architectural features found in this large sample of churches.

METHODOLOGY

Method Summary. The main research hypothesis is that the speech intelligibility in churches could be measured and related with objective room acoustics measures and architectural features. The among-room variations of speech intelligibility scores can be viewed as differences that result from the architectural and objective acoustical proprieties of the churches that experience shows actually exist.

The study consisted of two parts both regarding analyses in (almost) non occupied churches. The first part was to gather objective results of the main room acoustics measures. The second part was to gather subjective evaluations from listeners, using live speech, of the qualities of the churches concerning speech intelligibility.

There are certain limitations using this type of methodology for evaluations. The acoustical response of the church changes when it is fully occupied and the character of the sound heard during a religious service is likely different. Nevertheless this methodology gives a normalized sound environment that could be easily compared among churches.

Sample of Churches Used. This study reports on acoustical field measurements done between June 1993 and January 1996 in a major survey of 36 Roman Catholic churches in Portugal that were built between the 6th century and the 1960's. The churches are a sample of 14 centuries of church building in Portugal. Portugal is one of the oldest European countries and played a prominent role in some of the most significant events in worldwide history. It presents an almost perfect location to trace the history of Catholic church buildings in the world. Portuguese churches can be considered a representative example of Catholic churches in the world.

The churches were selected to represent the main architectural styles found throughout Portugal and to represent the evolution of church construction in Portugal. For more uniformity of the sample, only churches with a volume of less than 19000 m^3 were selected for the study.

Subjective acoustical evaluations were held in churches grouped by large periods of history: 12 *Visigothic* or *Romanesque* churches (6th-11th cent.), 11 *Gothic* or *Manueline* churches 13th-16th cent.), 9 *Renaissance* or *Baroque* churches (16th-18th cent.) and 4 *Neoclassic* or *Contemporary* churches (18th-20th cent.). The main architectural features of these churches are displayed in Table 1.

Table 1 - Simple architectural statistics for all 36 churches tested.

ARCHITECTURAL FEATURE	MINIMUM	MEDIAN	MEAN	MAXIMUM
Volume (m^3)	299	3829	5809	18674
Area (m^2)	56	424	448	1031
Maximum height (m)	6	14	15	39
Maximum length (m)	13	31	34	62
Width nave (m)	5	11	12	26

Measurement Method for Objective Measures. Six objective room acoustics parameters were calculated in each church using the Impulse Response Method (a sound source generates sound within the room and a receiving section acquires the sound pressure signal after the sound source ceases emit). They are:

RT Reverberation Time using the integrated impulse-response method. *RT30* (from -5 to -35 dB);

EDT Early Decay Time. *EDT10* (from 0 to -10 dB);

C80 Early to Late Sound Index or Clarity with a time window of 80 ms. $C80 = 10 \log E(0,80)/E(80,\infty)$;

D Early to Total Energy Ratio (Early Energy Fraction, Definition or *Deutlichkeit*) with a time window of 50 ms. $D = E(0,50)/E(0,\infty)$;

TS Center Time (point in time where the energy received before this point is equal to the energy received after this point);

L Loudness, Total Sound Level or Overall Level (measure of the room's ability to amplify sound from the source position). This measure is also denoted as *G* in the literature.

The method used is based on the integrated impulse-response method. A limited-bandwidth noise-burst is generated and transmitted into the church by a loudspeaker via an amplifier. The response of the room to the noise-burst (the *impulse response*) is then sampled from the RMS detector output of the sound level meter (time constant 5 ms). A loudspeaker emitting short pulses-noise bursts in 3/2 octave frequency bands (to ensure that the received noise-burst is of 1/1 octave bandwidth) was used as sound source. The receiving section consisted of one 1/2" microphone and a sound level meter with an octave filter set. All the procedure was controlled by a specific software. In each church, two sound source locations were used for the loudspeaker (in front of the altar and in the center of the main floor). The sound source was positioned at 0.8 m above the floor and making a 45° angle with the horizontal plane. Each measurement was calculated from an ensemble of 3 or 4 pulse responses in each position. Five receiver positions were, in average, used depending on the width of the church. The microphone, at each location, was placed at 1.30 m above the floor. In total, near 8000 values were determined (all combinations of the 6 octave-frequency bands, 125 to 4k Hz, and source-receiver locations).

Speech intelligibility was objectively quantified by the calculation of the Rapid Speech Transmission Index (*RASTI*) which was hypothesized that can be related to the scores of people taking live speech intelligibility tests. This method is based on the measurement of the reduction in signal modulation between the speaker and listener positions. A transmitter generates a special test signal (pink noise in the 500 and 2000 Hz octave bands) to mimic the long-term speech spectrum. An omnidirectional 1/2" diameter microphone receives the signal. The signal is transmitted to the *RASTI* receiver unit, which analyzes the signal and calculates the *RASTI* value. The receiver and transmitter are independent units (not synchronized) because the signal is repetitive. In each church the transmitter location was in front of the altar, 1.65 m above the floor. From 4 to 17 positions were used for the receiver depending on the length of the church (on average, 8 positions were used). In each receiver position three or four measurements were taken and then averaged together to give the *RASTI* value at that location.

The equipment used was sound level meter B&K-2231; filter set B&K-1625; module *Room Acoustics* B&K-BZ7109; sound source B&K-4224; microphone 1/2" diameter B&K; computer *Compaq* LTE with software *Room Acoustics* B&K-VP7155; speech transmission meter B&K-3361 and microphone 1/2" diameter B&K-4129.

Listeners and Sound Source. A group of 12 college students and 3 of their professors from the School of Music and the Performing Arts (Polytechnic Institute of Porto) was chosen. to judge the quality of speech throughout the churches. A young theater student from the same school, was used as a speaker. In each church he read a different list of 100 words within the same sentence: "This is (word)" (in Portuguese). The sentences were said with similar loudness and rhythm. The list of 100 words used in each church was chosen from an innovative global 400-word list that represents the Portuguese language.

In each church the listeners were seated in two similar locations named *Position A (*right hand seats of the center of the longitudinal axis of the main floor) and *Position B* (central seats at the rear main floor). Then they listed the words understood on a questionnaire sheet. The scores from the questionnaires were analyzed using the *SYSTAT®* software.

To qualify their answers, all members of this group of listeners performed audiometric tests to evaluate their hearing capabilities. Audiograms from 125 Hz to 8 kHz and according to ISO R389 and ANSI S3.6 were performed giving results judged normal for all the members of the listeners' group.

Only words with 1 to 4 syllables were used, but the 2 and 3-syllable words were predominant (96%) not to allow listeners to decipher the word by understanding only the sound of some syllables.

RESULTS

Relationships between Speech Intelligibility and Architectural Features. This chapter presents the results concerning the relationships between *Speech Intelligibility* and the architectural features of the churches. In this chapter all relationships are done with the averaged *Speech Intelligibility* data for each church (36 data points = 36 churches).

Table 2 presents the best simple models with the squared values for the correlation coefficients between *Speech Intelligibility* and each of the thirteen architectural parameters. The best relationship exists between *Speech Intelligibility* and *Height Nave* ($R^2 = 0.52$) presenting the importance that the nave height has regarding the *Speech Intelligibility*. However, other important relationship was found between *Speech Intelligibility* and *Height total average* ($R^2 = 0.50$). The Figure 1 shows the best fitting simple models, presented in Table 2.

To find a better model to explain the relationships between speech intelligibility and architectural features a general linear model was calculated. The operational procedure was to use the *forward* or the *backward* stepwise modeling with an α-to-enter/remove) equal to 0.15. The accuracy of the model is judged primarily by its R^2 which represents the percentage of variance explained and secondarily by the standard error of the estimate which represents the magnitude of differences between estimated and observed values. The general linear model is presented in Table 3 showing that the speech intelligibility is reasonably connected to the architectural features of the churches.

Table 2. Best simple models between *Speech Intelligibility* (in %) and the 13 Architectural Parameters.

EQUATIONS	R^2 (variance explained)
Speech = 150.347 - 8.410 log VTOT	0.494
Speech = 147.674 - 8.295 log VNV	0.493
Speech = 93.989 - 0.028 ATOT	0.405
Speech = 135.574 - 9.747 log ANV	0.371
Speech = 106.285 - 0.822 LMAX + 0.002 $LMAX^2$	0.419
Speech = 102.532 - 0.853 LNV	0.406
Speech = 110.453 - 2.563 HMAX + 0.034 $HMAX^2$	0.463
Speech = 104.587 - 1.893 HNV	**0.523**
Speech = 115.774 - 14.235 log WNV	0.237
Speech = 118.641 - 15.165 log WAVG	0.277
Speech = 110.154 - 2.914 VTAT + 0.031 $VTAT^2$	0.501
Speech = 91.729 - 0.070 ABST	0.329
Speech = 73.214 + 129.767 CABS	0.039

ABST-Total Absorption (m^2); ANV-Area Nave (m^2); ATOT-Area Total (m^2); CABS-Absorption Coef. α (avg value for all surfaces); HMAX-Height Maximum (m); HNV-Height Nave (m); LMAX-Length Maximum (m); LNV-Length Nave (m); VNV-Volume Nave (m^3); VTAT-Height Total avg (m) (=Volume total/Area total); VTOT-Volume Total (m^3); WAVG-Width avg (m);WNV-Width Nave (m). *Total* stands for the entire church including lateral chapels and main altar; *Nave* stands for the entire church excluding lateral chapels and main altar.

Table 3. Relationship between *speech intelligibility* (in %) and architectural parameters (general linear model).

GENERAL LINEAR MODEL EQUATION	St. Error of Estimate	R^2
Speech = 104.5 - 8.9×10^{-3} VTOT + 6.4×10^{-3} VNV + 0.091 ATOT - 1.192 LNV + 1.441 HMAX- 1.915 HNV - 1.830 WNV + 188.7 CABS	7.8	0.72

ATOT-Area Total (m^2); CABS-Absorption Coef. α (avg value for all surfaces); HMAX-Height Maximum (m); HNV-Height Nave (m); LNV-Length Nave (m); VNV-Volume Nave (m^3); VTOT-Volume Total (m^3); WNV-Width Nave (m). *Total* stands for the entire church including lateral chapels and main altar; *Nave* stands for the entire church excluding lateral chapels and main altar.

Relationships between Speech Intelligibility and Objective Acoustical Parameters.
AVERAGING METHOD. The following analyses were done with averaged data for each church. Seven frequency averaging methods were tested using the average of 2 to 6 octave frequency-bands to obtain a single-number for each objective room acoustic parameter and for each church. These options were named M1 to M7 and are explained in Table 4. Regression analyses were performed with each of these seven averaging options to check for their influence in the results.

Table 4 - Seven options of frequency averaging methods (*Mi*).

CODE	DEFINITION	RANGE
M1	Average of all 6 frequency bands	125 to 4000 Hz octave bands
M2	Average of the 2 highest frequency bands	2000 and 4000 Hz octave bands
M3	Average of the 4 lowest frequency bands	125 to 1000 Hz octave bands
M4	Average of the 4 highest frequency bands	500 to 4000 Hz octave bands
M5	Average of 4 medium frequency bands	250 to 2000 Hz octave bands
M6	Average of 3 medium frequency bands	500, 1000 and 2000 Hz octave bands
M7	Average of 2 medium frequency bands	500 and 1000 Hz octave bands

SIMPLE MODELS. Using each of the frequency averaging options *Mi* shown in Table 4, linear and non linear models were tested for the relationships between *speech intelligibility* and the six objective room acoustic parameters. Table 5 presents the equations for the best models found and for each objective acoustical parameter. The Figure 2 presents the best relationships found.

It is shown that the variance of the *speech intelligibility* can be fairly explained with just one of the six objective room acoustic parameters ($R^2 \approx 0.67$). For *D* and *L* the percentage of variance explained by just one objective room acoustic parameter is not very significant ($R^2 < 0.47$).

The relationship *Speech*/RT ($R^2 = 0.674$) confirms that RT has influence in the *speech intelligibility*. However, using TS the R^2 is very similar (0.665) making this objective room acoustic measure almost as suited as RT to fairly estimate the speech intelligibility. The relationship *Speech*/L ($R^2 = 0.33$) does not fulfill the reasonable expectations regarding their connection.

Table 5. Most significant relationships between *speech intelligibility* (in %) and each of the six objective acoustical parameters (with the frequency averaging method *Mi*).

METHOD *Mi*	EQUATIONS (simple models)	R^2 (variance explained)
M6	*Speech* = 118.104 - 14.519 RT + 0.907 RT^2	**0.674**
M7	*Speech* = 117.157 - 13.966 EDT + 0.860 EDT^2	0.660
M4	*Speech* = 95.162 + 2.255 C80 - 0.331 $C80^2$	0.644
M7	*Speech* = 53.381 + 197.581 D - 215.340 D^2	0.465
M7	*Speech* = 116.937 - 0.181 TS + 1.4 x 10^{-4} TS^2	0.665
M2	*Speech* = 38.653 + 5.051 L - 0.116 L^2	0.328

GENERAL LINEAR MODELS. To find a better model to explain the relationship between *speech intelligibility* and objective acoustical parameters, a general linear model was calculated. The operational procedure was identical as the one described before.

The parameter *C80* appeared as variable in all seven general linear models tested (one for each *Mi*) indicating that this measure can be very important in predicting the speech intelligibility of churches. *D* performs almost as well as *C80* in that function. *L* however, does not appear in any of the models, perhaps revealing that it is not a significant measure in predicting speech intelligibility in churches. The two best general linear models with their frequency averaging option (*Mi*) are presented in Table 6.

Table 6. Best two relationships between *speech intelligibility* (in %) and objective acoustical parameters (general linear models with the frequency averaging method *Mi* used).

METHOD	EQUATIONS (general linear models)	St. Error of Estimate	R^2 (variance explained)
M1	*Speech* = 140.34 - 40.957 RT + 43.354 EDT + 8.924 C80 - 136.724 D	6.9	**0.745**
M3	*Speech* = 161.41 + 10.190 C80 - 200.874 D	7.2	0.705

Relationship between Speech Intelligibility and the Objective Parameter RASTI. Similarly as presented above, statistical models were calculated to quantify relationships between *Speech Intelligibility* and *RASTI* values. To obtain relationships using the *RASTI* values, there was a need for a corresponding single *RASTI* value representative of each church. Two options were tested:

RASTI.avg - Average of all positions in each church;

RASTI.ndf - Average of all positions Not in the Direct Field of the sound source (that is, excluding positions < 5 m from sound source or not in the main volume of the church).

As shown in Table 7, *RASTI.ndf* appears as the best of the two parameters to predict speech intelligibility (R^2 = 0.503). However, this fairly small R^2 reveals that there is no significant relationship between church averaged *RASTI* and speech intelligibility values. It is presumed that the results would be better if there were a direct spatial comparison (relating measure point with measure point) and not just church averaged values to compare. Figure 3 presents the regression model between *RASTI.ndf* and *speech intelligibility*.

Figure 4 shows, for each church, the ratios between averaged *speech intelligibility* and *RASTI* values where it is seen that the *speech intelligibility* values are, in general, the double of the corresponding mean *RASTI* (from x1.3 to x2.5). Table 8 proposes a simple new scale to tie *RASTI* values to subjective speech intelligibility.

Table 7. Best models and R^2 between *Speech Intelligibility* (in %) and the objective parameter *RASTI* (in %).

Objective Parameter	EQUATION	R^2 (variance explained)
RASTI.avg	*Speech* = -52.54 + 4.951 RASTI.avg - 0.041 RASTI.avg^2	0.480
RASTI.ndf	*Speech* = -50.55 + 5.315 RASTI.ndf - 0.048 RASTI.ndf^2	**0.503**

Table 8. Proposed table to convert from RASTI values to a subjective intelligibility scale (* proposed).

RASTI (in %)	*Speech* (in %)*	Subjective speech intelligibility scale
0 - 30	0 - 60	*bad*
30 - 45	60 - 75	*poor*
45 - 60	75 - 85	*fair*
60 - 75	85 - 95	*good*
75 - 100	95 - 100	*excellent*

CONCLUSIONS

Statistically significant relationships between *speech intelligibility* and objective criteria can be found in churches.

Architectural features that are important to defining the *speech intelligibility* in churches were identified. *Nave Height* was found as the most important of these, giving the best fit. In general, 8 of the 13 architectural parameters tested can be used in a general linear model to explain 72% of the *speech intelligibility* variance.

The best fit between *speech intelligibility* and objective acoustical parameters was for *Speech/RT* (R^2 = 0.67) stating that *RT* can be a reasonable predictor of the *speech intelligibility*. The relationship *Speech/TS* (R^2 = 0.67) also confirms the suitability of *TS* to fairly and objectively represent the *speech intelligibility*. The relationship *Speech/L* (R^2 = 0.33) does not fulfill the expectations regarding their connection.

Some of the six objective acoustical parameters tested can be used in a general linear model to explain 75% of the variance of the *speech intelligibility*. This best general model was found to be with the frequency averaging option M1 (average of all 6 frequency bands from 125 to 4000 Hz). C80 is the only objective acoustical parameter that appears in all the general linear models found, indicating that it is also an meaningful measure to predict *speech intelligibility*.

The relationship between *speech intelligibility* and *RASTI* was tested using two methods to find a representative single number: *RASTI.avg* (average of all positions in each church) and *RASTI.ndf* (average of all positions Not in the Direct Field of the sound source). *RASTI.ndf* appears as the best to predict *speech intelligibility* (R^2 = 0.503). However, this fairly small R^2 reveals that there is no significant relationship between church averaged *RASTI* and *speech intelligibility* values. It is presumed that the results would be better if there were a direct spatial comparison (relating measure point with measure point) and not just church averaged values to compare. It was also found that the *speech intelligibility* values were, in general, the double of the corresponding mean *RASTI*. A simple new table was proposed to tie *RASTI* values to a subjective speech intelligibility scale.

REFERENCES

Carvalho, António P. O. "Influence of Architectural Features and Styles on Various Acoustical Measures in Churches", Ph.D. Diss., (U of Florida, Gainesville, 1994);

Carvalho, António P. O., A. Morgado and L. Henrique "Analysis of Subjective Acoustic Measures and Speech Intelligibility in Portuguese churches". Paper presented at the 131st Meeting of Acoust. Soc. Am., Indianapolis, (May 1996);

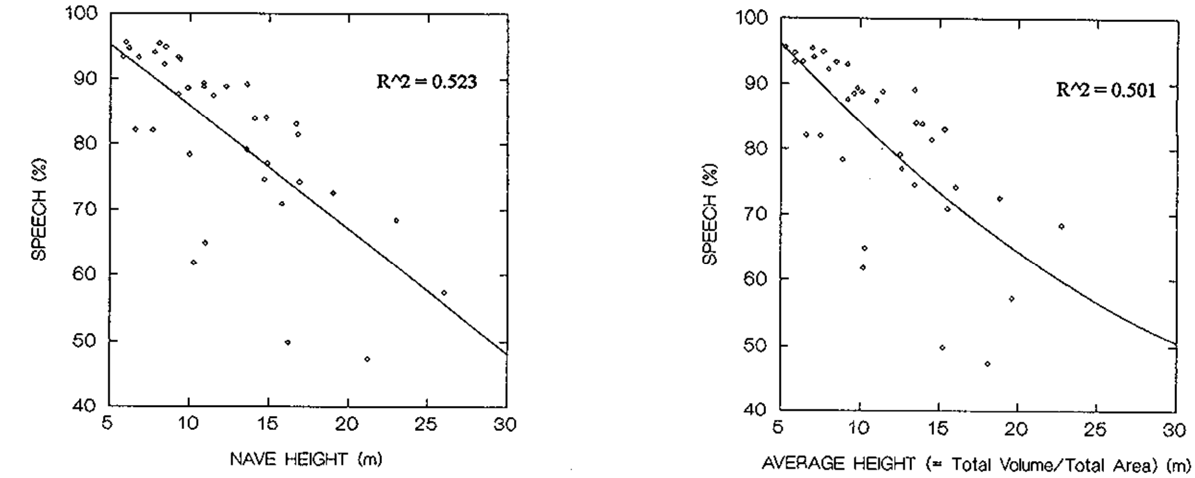

Figure 1 - Best regression models relating *speech intelligibility* vs. architectural parameters mean values (36 points = 36 churches).

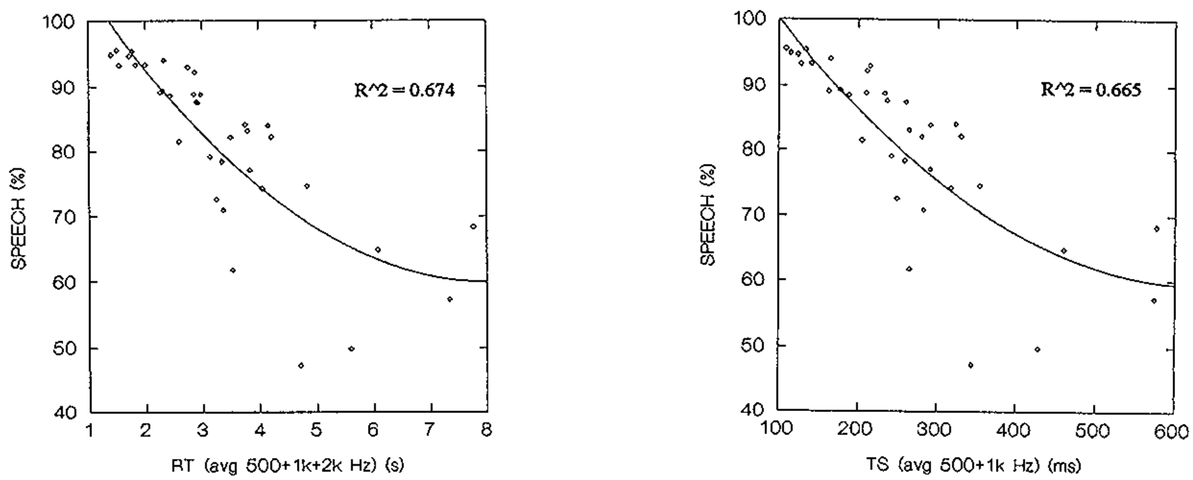

Figure 2 - Best regression models relating *speech intelligibility* vs. objective acoustical parameters mean values (36 points = 36 churches).

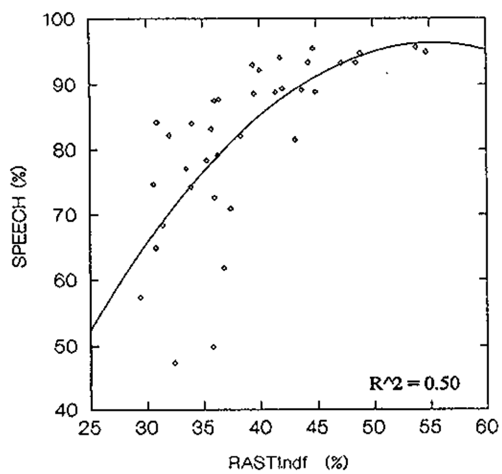

Figure 3 - *Speech intelligibility* vs. *RASTI.ndf* (not in the direct field) mean values with regression model (36 points = 36 churches).

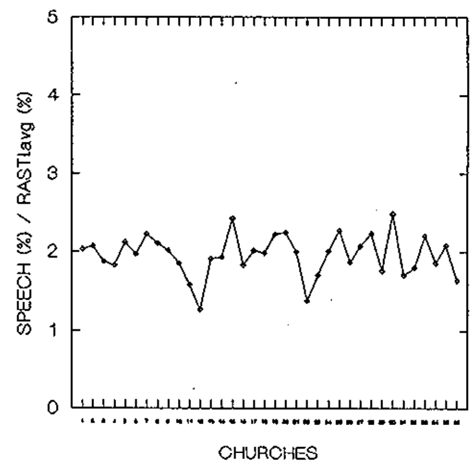

Figure 4 - Ratios between church averaged *speech intelligibility* and *RASTI* values.