NATIONAL TRANSPORTATION SAFETY BOARD
OFFICE OF AVIATION SAFETY (AS-50)
Washington, D.C. 20594

July 19, 2000

SPEECH EXAMINATION STUDY

A. ACCIDENT: DCA-00-MA-006

Operator: EgyptAir Flight MS-990
Location: Off Quonset Point, RI
Date: October 31, 1999
Time: 0148 Eastern Standard Time
Airplane: Boeing 767-300 SU-GAP

B. SPEECH EXAMINATION GROUP

Malcolm Brenner (Chair),
National Transportation Safety Board (NTSB)

With assistance from:

Vernon S. Ellingstad, Jim Cash, Rob Molloy, Miriam Kloeppel
National Transportation Safety Board (NTSB)

Captain Shaker Kelada, Engineer Waguih Sobhi Tadros
EgyptAir

C. SUMMARY

About 0150 eastern standard time (EST), on October 31, 1999, a Boeing 767-366ER, SU-GAP, operated by EgyptAir, as flight 990, crashed into the Atlantic Ocean about 60 miles south of Nantucket, MA. EgyptAir flight 990 was being operated under the provisions of Egyptian Civil Aviation Regulations Part 121 and United States Title 14 Code of Federal Regulations Part 129 as a scheduled, international flight from John F. Kennedy Airport (JFK), New York, New York to Cairo International Airport in Cairo, Egypt. The flight departed JFK about 0122 EST, with 4 flightcrew members, 10 flight attendants, and 203 passengers on board. There were no survivors. The airplane was destroyed by impact forces.

D. DETAILS OF INVESTIGATION

Following the work of the Cockpit Voice Recorder (CVR) Group, which prepared a transcript of the conversation on the CVR, a speech examination study was conducted to examine scientific approaches to speech analysis that could verify or expand on the information of the transcript.[1] The transcript indicated that the captain left the cockpit for a toilet break about 1:48:18 (EST), and that, following this, there were 23 statements during an emergency period that included the airplane departing from level flight (until the CVR stopped recording at 1:50:38). The CVR Group identified the speakers of these 23 statements as an unidentifiable speaker (1:48:30), First Officer Batouti (1:48:39; 1:49:48; 1:49:57; 1:49:58; 1:50:00; 1:50:01; 1:50:02; 1:50:04;1:50:05; 1:50:07; 1:50:08; 1:50:29), and Captain Habashi (1:50:06; 1:50:08; 1:50:15; 1:50:24; 1:50:26; 1:50:28; 1:50:31; 1:50:32; 1:50:34; 1:50:36).

Because of the importance of this material to the investigation, this study examined linguistic evidence and laboratory measures of speech that might provide:

- any evidence that would help identify the speaker of the statement at 1:48:30 as well as what was said;
- any evidence that would help identify the speakers of the remaining statements made after 1:48:30, including evidence on how many speakers were present in the cockpit;
- any evidence that would help determine whether speakers were expending a high physical effort at any time during the emergency period; and,
- any additional speech evidence that might assist the investigation.

The observations of the study follow.

Linguistic Evidence

The linguistic content of what was said can provide evidence on the identity of the speakers.

The most direct linguistic evidence was provided in a statement "Gamil, what's happening?" (0150:08) in which the speaker identified another crewmember by name. First Officer Batouti was the only crewmember with the personal name "Gamil." It appears that the speaker was acquainted with Batouti, recognized him in the cockpit, and believed that Batouti could help explain what was happening.

[1] For additional speech evidence, see the "Group Chairman's Factual Report Sound Spectrum Study, Cockpit Voice Recorder."

Further, the speaker provided evidence on his own identity by addressing Batouti by a personal name. According to representatives of EgyptAir, only Captain Habashi or Captain Hatem would have had sufficient seniority and status among the available crewmembers to have addressed First Officer Batouti by his personal name.[2]

With regard to the unknown statement at 1:48:30, linguistic evidence suggests it was isolated rather than part of a conversation. It was preceded by 19.2 seconds without speech, and followed by 9.2 seconds without speech until the start of a series of 11 repetitions of "tawkalt." There was no evidence that it formed part of an interchange between two persons.

Laboratory measures

Computer-based measurements for this study were made of three aspects of speech—fundamental frequency,[3] formant dispersion,[4] and speech duration[5]—to provide evidence related to speaker identification and psychological stress.

Extraction of the speech sample

The CVR provided an audio record of the last 31.4 minutes of the flight as recorded at a rate of 1 7/8 inches per second on a Fairchild Model A-100 analog tape recorder located in the tail of the airplane. All speech samples analyzed in the study were captured through the cockpit area microphone located in the overhead panel. Additional audio information was captured by the first officer's hot microphone (which was worn by First Officer Adel during the first part of the tape and then rested on the right side of the cockpit when First Officer Batouti assumed pilot duties). This hot mike information was reviewed but did not provide useful speech information during the emergency period.[6]

The CVR tape was digitized at a rate of 20 KHz., and the digital version was reviewed by four EgyptAir representatives who spoke Arabic as a native language and were familiar with the crewmembers.[7] They identified speakers for each recorded

---

[2] According to EgyptAir representatives, Captain Hatem smoked heavily and was responsible for the "sounds similar to coughing" noted 6 times in the CVR transcript.

[3] Fundamental frequency is the rate at which the vocal cords of the larynx open and close during speech releasing puffs of air. A fundamental frequency of 150 Hz. indicates that the vocal cords open and close 150 times per second. Listeners normally perceive the fundamental frequency as the pitch of the speaker's voice.

[4] The formants are frequencies at which the vocal tract above the larynx, acting as a filter due to its natural modes of vibration, will allow maximum energy to pass from the sound produced by the vocal cords. The formant frequencies determine many aspects of perceived speech. Formant dispersion refers to the relative spacing between successive formants.

[5] Speech duration refers to how long the speech lasts.

[6] The captain's hot microphone audio channel was difficult or impossible to read.

[7] The four representatives were Captain Shaker Kelada, Engineer Waguih Sobhi Tadros, Engineer Fatma Ismael, and Engineer Yousaf Abdallah.

utterance on the tape prior to 1:48:18 (where utterance was defined on a case-by-case basis by the group members as a discrete segment of unbroken speech). For an utterance to be used as data, the group members had to agree unanimously on the identity of the speaker and the utterance had to provide a sufficient signal-to-noise ratio to permit computer scoring.

All utterances were analyzed through the Entropic Signal Processing System for computer analysis (Entropic Research Laboratory, Inc., Washington, D.C.). The program detected voicing[8] at regular points within the waveform, and, for voiced samples, selected fundamental frequency and formant estimates from candidates proposed by solving for the roots of the linear predictor polynomial computed periodically from the waveform.

Speech samples and analysis parameters were portrayed on a Waves digital video display and samples could also be heard through speakers. When the start and stop times of a desired utterance were captured, these times were recorded and provided a measurement of segment duration.

Prior to 1:48:18, 6 EgyptAir pilots visited the cockpit and made statements that were recorded and analyzed. Numerous speech samples were obtained for Captain Habashi (51 utterances, totaling 162.8 seconds of speech), First Officer Batouti (27 utterances, totaling 46.0 seconds), and First Officer Adel (24 utterances, totaling 40.1 seconds). Additional speech samples were identified for Captain Hatem (10 utterances, totaling 23.0 seconds of speech), while limited speech samples were obtained for First Officers Hisham (5 utterances, totaling 6.1 seconds) and Aiad (2 utterances, totaling 3.0 seconds).[9]

All utterances made after 1:48:18 were analyzed individually. Due in part to loud background noise in the cockpit during this period, only 15 of the 23 (65%) utterances in the CVR transcript provided sufficient signal-to-noise properties to be analyzed by computer for fundamental frequency and formant information.

Fundamental Frequency

Fundamental frequency varies characteristically among speakers and often conveys information on the speaker's psychological stress [Refs. 1, 2, 3, 4, 7, 8, 10, 11].

Figure 1 summarizes the fundamental frequency information obtained from the present data. The first 6 entries summarize the distribution of fundamental frequency scores observed for each EgyptAir pilot during the routine portion of the flight. The square represents the mean observation, while the bars display two standard deviations

[8] Voicing refers to sounds produced by a periodic motion of the vocal cords, contrasted with unvoiced sounds which are produced by blowing air through a narrow opening (such as between the tongue and the roof of the mouth). Fundamental frequency and formant measures apply only to voiced samples.
[9] A female flight attendant spoke briefly on the CVR during this period but her speech was not sufficiently louder than background noise to permit computer analysis.

around the mean in each direction. The 6 pilots varied significantly in fundamental frequency, as confirmed by an analysis of variance statistical test [speaker effect: $F(5/113) = 13.2$, $p<.001$, of mean differences in fundamental frequency between the six speakers]. The absolute values of fundamental frequency for individual speakers were:

Speaker	Average fundamental frequency during routine portion of flight
First Officer Adel	194.8 Hz.
First Officer Aiad	193.0 Hz.
First Officer Batouti	162.6 Hz.
Captain Habashi	203.1 Hz.
Captain Hatem	169.3 Hz.
First Officer Hisham	148.2 Hz.

The remaining entries in Figure 1 plot, from left to right, the fundamental frequency for each successive statement made during the emergency.

The first statement, marked "*," was the unknown statement at 1:48:30. It displayed a fundamental frequency of 163.0 Hz..

The remaining statements in Figure 1, all made after the airplane departed from normal flight, displayed an interesting difference. Some statements showed a fundamental frequency value that appears to be within ranges shown in routine speech by the 6 pilots, including all statements of "tawkalt" and the statement "it's shut." By contrast, the remaining statements showed a fundamental frequency value that was high compared to the routine speech of almost all 6 pilots.

Belan (Refs. 1, 4) provides guidelines to evaluate the degree of psychological stress that a speaker may experience based on an increase in fundamental frequency. According to Belan, an increase of about 30% in fundamental frequency, when compared to relaxed levels for the same speaker, characterizes a first stage of stress during which attention is mobilized and performance improves. A further increase, of 50-150%, characterizes a second stage of stress during which a speaker experiences strain and performance may become hasty and abbreviated but not display gross mistakes. A final stage, characterized by an increase of 100-200%, represents panic during which the pilot can not think straight. According to these guidelines, the speaker(s) of "tawkalt" and "it's shut" was expressing no more than a first stage of stress, while the speaker of "pull with me" was expressing a higher stage of stress in response to the emergency nature of the situation.

Formant Dispersion[10]

Formant dispersion, according to recent scientific literature, may vary inversely with the height of the speaker because of individual differences in the size of the vocal tract [Refs. 5, 6, 9]. If true, this speech property would provide a useful descriptive measure for use in speaker identification.

Formant dispersion was determined in this study for all 6 pilots by means of automated extraction of the 3 lowest speech formants (in Hertz) and, following Fitch [Ref. 5, p. 1216], calculation of the average differences separating successive formants.[11]

Figure 2 summarizes the relationship between the height of each of the 6 EgyptAir pilots (as reported on their most recent pilot medical examination) and the average formant dispersion measure obtained for them during the routine portion of the flight. As shown in Figure 2, there was a modest but not significant inverse relationship between the height of the speaker and the formant dispersion measure obtained ($r = -.635$, $n = 6$, n.s.). This correlation was significant when the data were restricted to those four pilots (Adel, Batouti, Habachi, Hatem) from whom longer speech samples were available (23.0 seconds of speech or more)($r = -.97$, $n = 4$, $p<.05$). Future research needs to determine the minimum length of speech sample necessary to produce an accurate formant dispersion measure. Future research is also needed to confirm a relationship between height and formant dispersion given the limited evidence available in the present study and previous literature.

Figure 3 summarizes the formant dispersion results obtained from the present data. The first 6 entries summarize the distribution of formant dispersion scores observed for each of the EgyptAir pilots during the routine portion of the flight. The square represents the mean for each pilot, while the bars display two standard deviations around the mean in each direction. The 6 pilots varied significantly in formant dispersion, as confirmed by an analysis of variance statistical test [speaker effect: $F (5/109) = 15.2$, $p<.001$, of differences in mean dispersion scores between the 6 speakers]. The absolute values of formant dispersion scores for individual speakers were:

Speaker	Average formant dispersion during routine portion of flight
First Officer Adel	975.2 Hz.
First Officer Aiad	957.1 Hz.

[10] Dr. Philip Lieberman, Linguistics Department, Brown University, Providence, RI, assisted the Safety Board by proposing this measure for speaker identification and consulting on various aspects of the scientific analysis.

[11] At the suggestion of Dr. Tecumseh Fitch, Program in Speech and Hearing Sciences and Department of Anthropology, Harvard University (personal communication), the measure was calculated using the lowest three formants although exploratory work suggested that the fourth formant also added meaningful information in the present data.

First Officer Batouti	933.4 Hz.
Captain Habashi	1040.9 Hz.
Captain Hatem	986.3 Hz.
First Officer Hisham	943.1 Hz.

The first statement among the remaining entries, marked "*," was the unknown statement at 1:48:30. It displayed a formant dispersion of 835.4 Hz. This value was at the lower extreme of the distribution shown by the pilots during routine conversation.

Several possible limitations might be reflected in this finding. The content of the "*" statement was unknown, and values of formant dispersion vary depending on the actual sounds being produced [Ref. 9]. It is possible that the unknown vowel sounds involved in this statement produced an artifactually low value of formant dispersion from one of the known speakers. This problem was especially relevant because of the short duration of the sample available for analysis (about 300 milliseconds) and the fact that the sample was characterized by only 1 data point.

The statement "it's shut" displayed a formant dispersion value of 908.8 Hz. This statement occurred well into the emergency period and there is presently no scientific literature on the effect of psychological stress on speech formant dispersion characteristics to assist at interpreting this finding.

Speech duration

Speech duration varies among speakers and often conveys information on the speaker's psychological stress [Refs. 3, 4, 7, 10, 11].

In the present study, the unknown statement at 1:48:30 was of too short duration to allow speech duration to be used as a reliable measure for speaker identification.[12] However, the statement "tawkalt", which was repeated 11 times as the airplane departed progressively from routine flight, provided a statement of known content on which useful measures of duration could be completed.

Figure 4 summarizes the speech duration information obtained from the 11 successive utterances of the phrase "tawkalt". The first utterance, made when the airplane was level, displayed the longest duration (1.02 seconds). The remaining utterances, made as the airplane continued into a progressively longer departure from level flight, varied in duration from 730-870 milliseconds but appeared to be relatively regular. Speech duration often becomes shorter with increasing psychological stress (i.e.

[12] Speaking duration measures (or speaking rate measures, which are simply the inverse of speaking duration) are generally not performed with utterances shorter than 5 syllables, unless the statement used in all test samples is identical [Refs. 4, 8]. The unknown statement consists of only 3 syllables of unknown content.

speaking rate becomes faster)[Refs. 3, 4, 5, 10]. However, the speaker of the phrase "tawkalt" did not appear to display a significant shortening in the speech duration.

Figures 5 summarizes data on the duration of the pause between successive utterances of "tawkalt". There was a pause of 67.5 seconds between the first and second utterances (not shown in Figure 6), a pause of 8.1 seconds between the second and third utterances (not shown in Figure 6), and pauses of between 510-700 milliseconds (as shown in Figure 6) between successive utterances. After the second utterance, the data suggested a rhythmic repetition of the phrase rather than an accelerating trend that might be expected with increases in psychological stress. The speaker displayed no clear evidence of increased psychological stress after the first or second utterances of the "tawkalt" phrase.

The speaker of the phrase "pull with me" showed changes in duration from the first utterance (630 milliseconds), to the second utterance (700 milliseconds), to the final utterance (730 milliseconds). This paralleled a changing trend of fundamental frequency values over these statements (Figure 1), with both effects suggesting a slight relaxation of psychological stress over successive statements.

Additional observations

An attempt was made to detect whether any additional human speech was present from 1:48:18 to the end of the tape. The tape was reviewed by a group of listeners in repeated segments, with computer analysis used to search for human voice-like harmonics where appropriate. No additional human speech was detected.

[redacted]

7-19-00

Malcolm Brenner, Ph.D.
Senior Human Performance Investigator

REFERENCES

1. Aircraft Accident Investigation Report [Japan], Japan Air Lines Co., Ltd., Boeing 747 SR-100, JA8119, Gunman Prefecture, Japan, August 12, 1985.

2. Bureau of Air Safety Investigation Accident Investigation Report, Mid-air collision between Cessna 172-N VH-HIZ and Piper PA38-112 VH-MHQ, near Tweed Heads, New South Wales, 20 May 1988 (BASI Report 881/1042)

3. Brenner, M., Doherty, E.T., Shipp, T. (1994). Speech measures indicating workload demand. *Aviation, Space, and Environmental Medicine, 65*:21-6.

4. Brenner, M., Mayer, D., Cash. J. [NTSB] (1996). Speech analysis in Russia. In *Methods and Metrics of Voice Communications.* Ed. B.G. Kanki, O.V. Prinzo. Washington, D.C.: Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, DOT/FAA/AM-96/10.

5. Fitch, W. T. (1997). Vocal tract length and formant frequency dispersion collelate with body size in rhesus macaques. *Journal of the Acoustical Society of America*, 102 *(2)*:1213-1222.

6. Fitch, W. T. (1999). Morphology and development of the human vocal tract: A study using magnetic resonance imaging. *Journal of the Acoustical Society of America*, *106 (3)*: 1511-1522.

7. Mayer, D.L., Brenner, M., Cash, J.R. Development of a speech analysis protocol for accident investigation. In *Methods and Metrics of Voice Communications.* Ed. B.G. Kanki, O.V. Prinzo. Washington, D.C.: Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, DOT/FAA/AM-96/10.

8. National Transportation Safety Board (1999). Uncontrolled descent and collision with terrain, USAir Flight 427, Boeing 737-300, N513AU, near Aliquippa, Pennsylvania, September 8, 1994. Aircraft Accident Report, NTSB/AAR-99/01, Washington, D.C., pp. 135-145.

9. Peterson, G.E., Barney, H.L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, *24 (2),* 175-184.

10. Ruiz, R., Legros, C., Guell, A. (1990). Voice analysis to predict the psychological or physical state of a speaker. *Aviation, Space, and Environmental Medicine, 61*:266-71.

11. Williams, C.E., Stevens, K.N. (1981). Vocal correlates of emotional states. In: Darby, J.K., ed. Speech evaluation in psychiatry. New York, NY: Grune & Stratton.

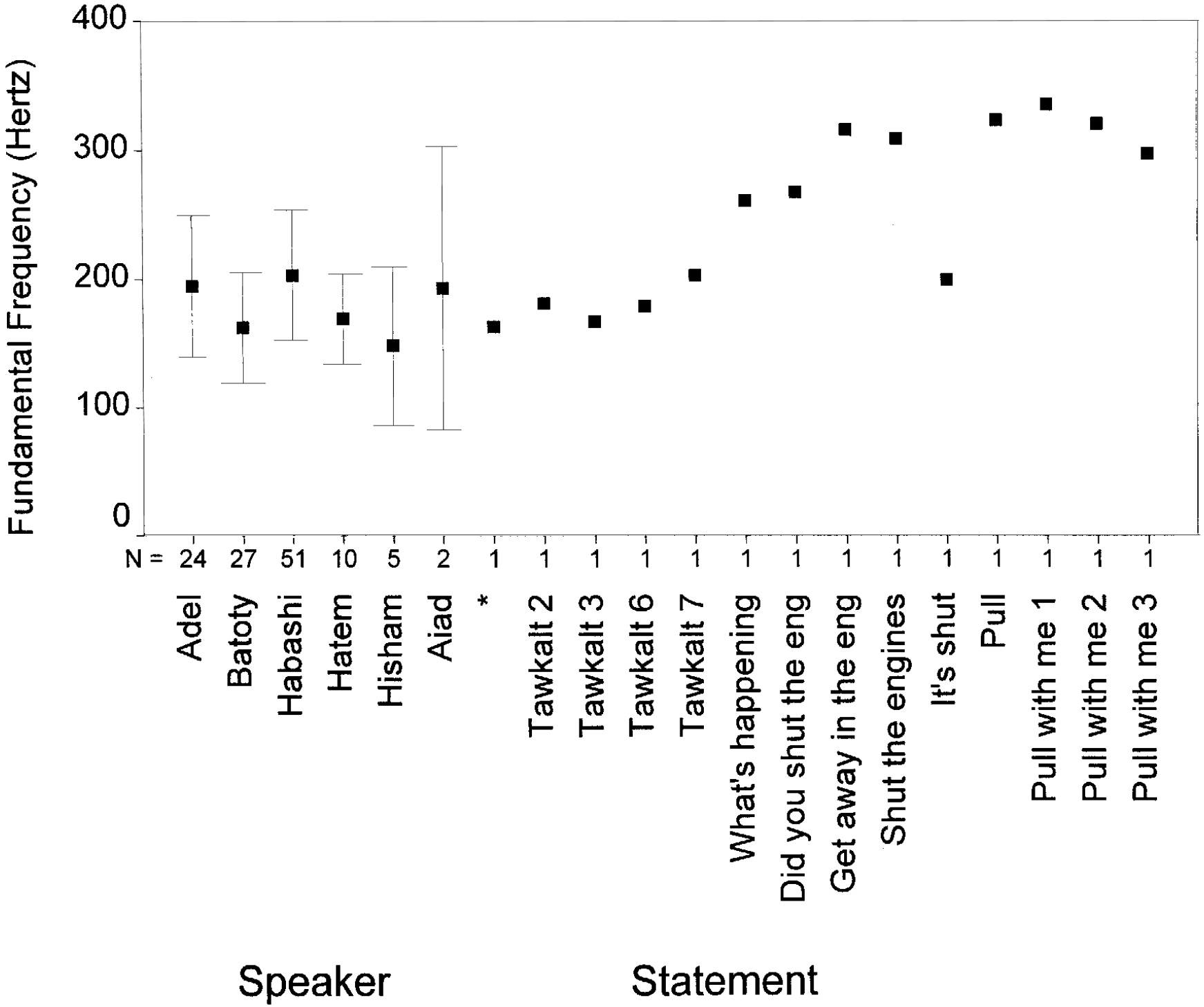
Fundamental Frequency (Hertz)
400
300
200
100
0
N = 24 27 51 10 5 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Adel
Batoty
Habashi
Hatem
Hisham
Aiad
*
Tawkalt 2
Tawkalt 3
Tawkalt 6
Tawkalt 7
What's happening
Did you shut the eng
Get away in the eng
Shut the engines
It's shut
Pull
Pull with me 1
Pull with me 2
Pull with me 3
Speaker
Statement

FIGURE 1

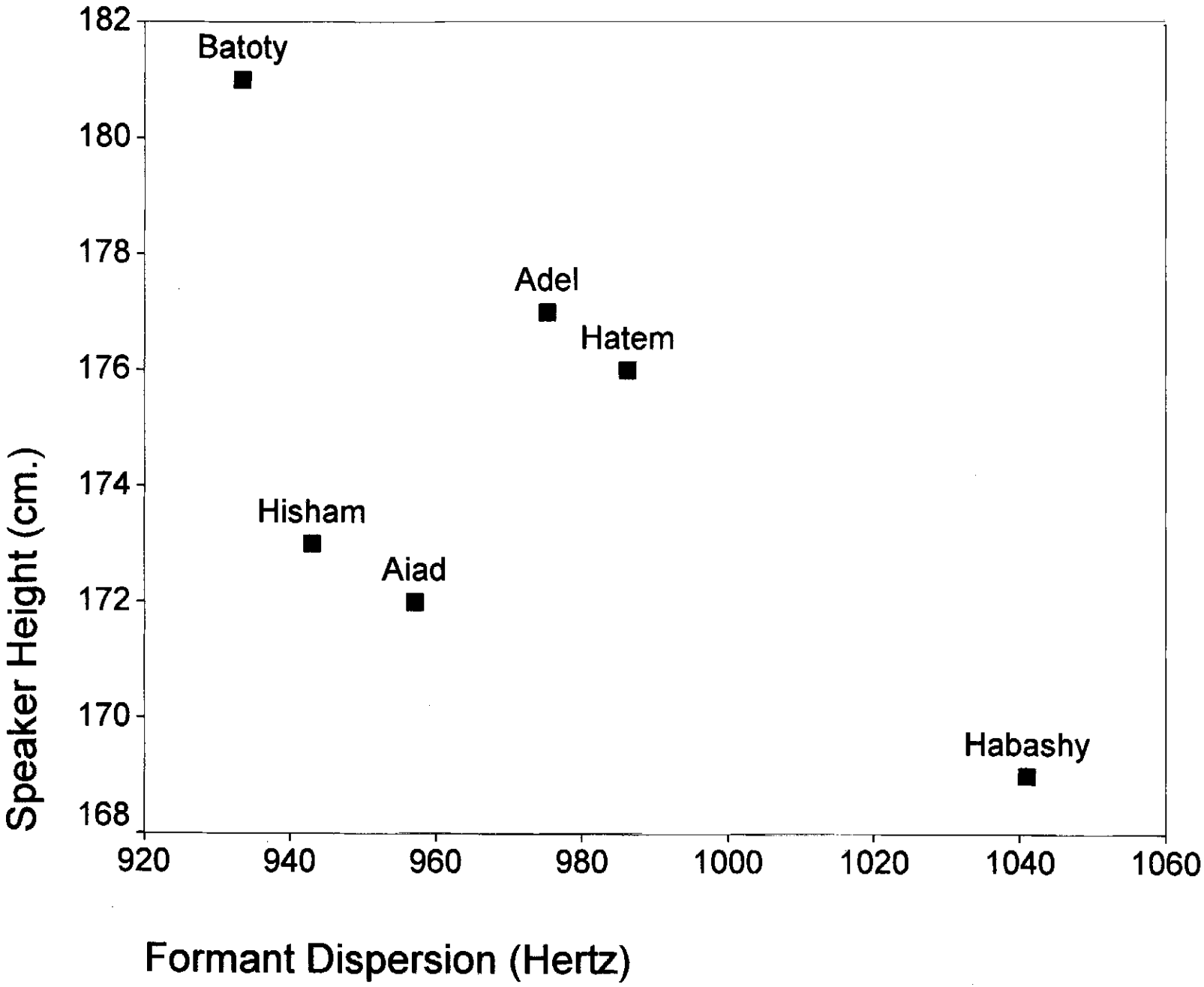
182
180
178
176
174
172
170
168
Batoty
Adel
Hatem
Hisham
Aiad
Habashy
Speaker Height (cm.)
920
940
960
980
1000
1020
1040
1060
Formant Dispersion (Hertz)

FIGURE 2

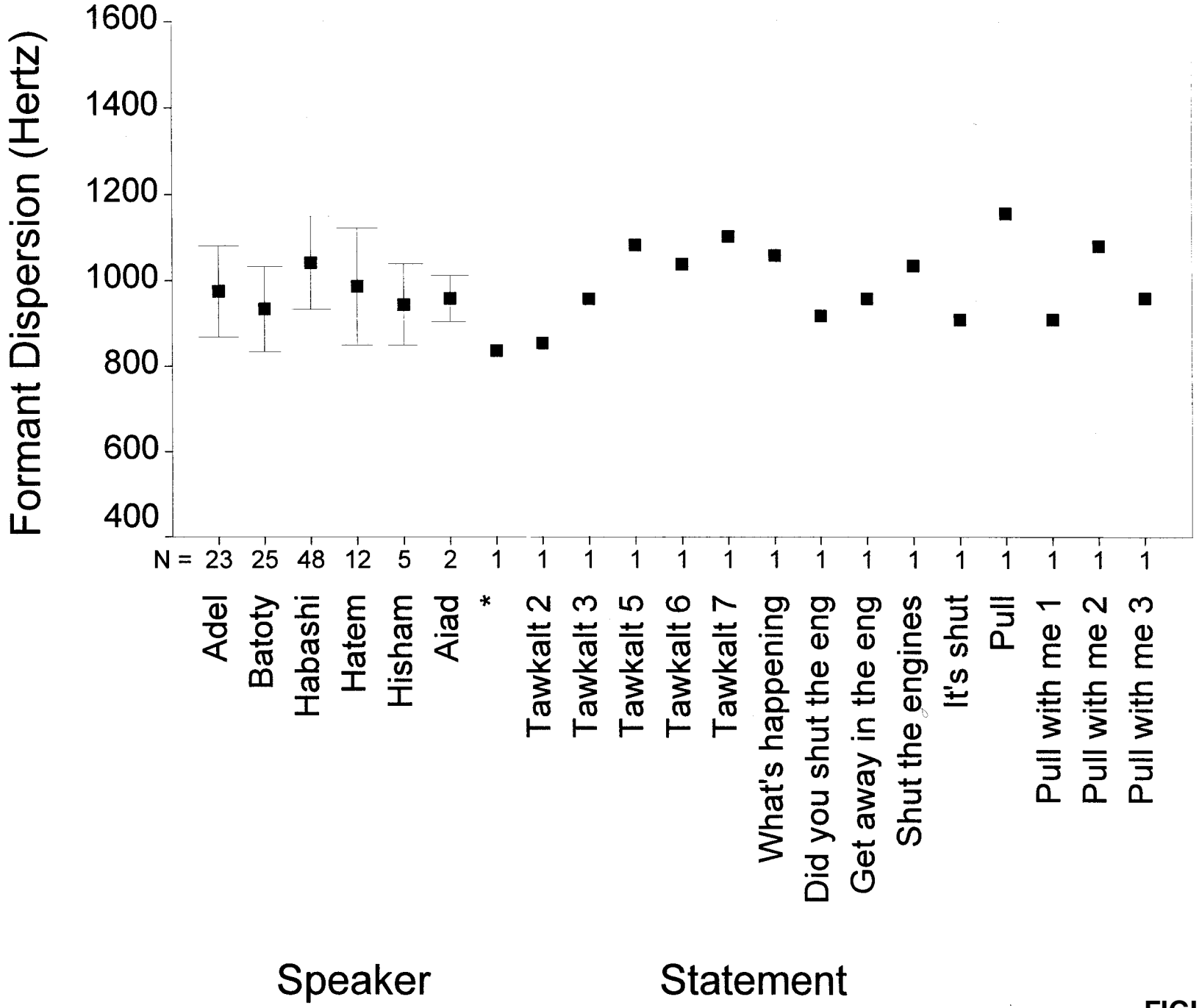
Formant Dispersion (Hertz)
1600
1400
1200
1000
800
600
400
N = 23 25 48 12 5 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Adel
Batoty
Habashi
Hatem
Hisham
Aiad
*
Tawkalt 2
Tawkalt 3
Tawkalt 5
Tawkalt 6
Tawkalt 7
What's happening
Did you shut the eng
Get away in the eng
Shut the engines
It's shut
Pull
Pull with me 1
Pull with me 2
Pull with me 3
Speaker
Statement

FIGURE 3

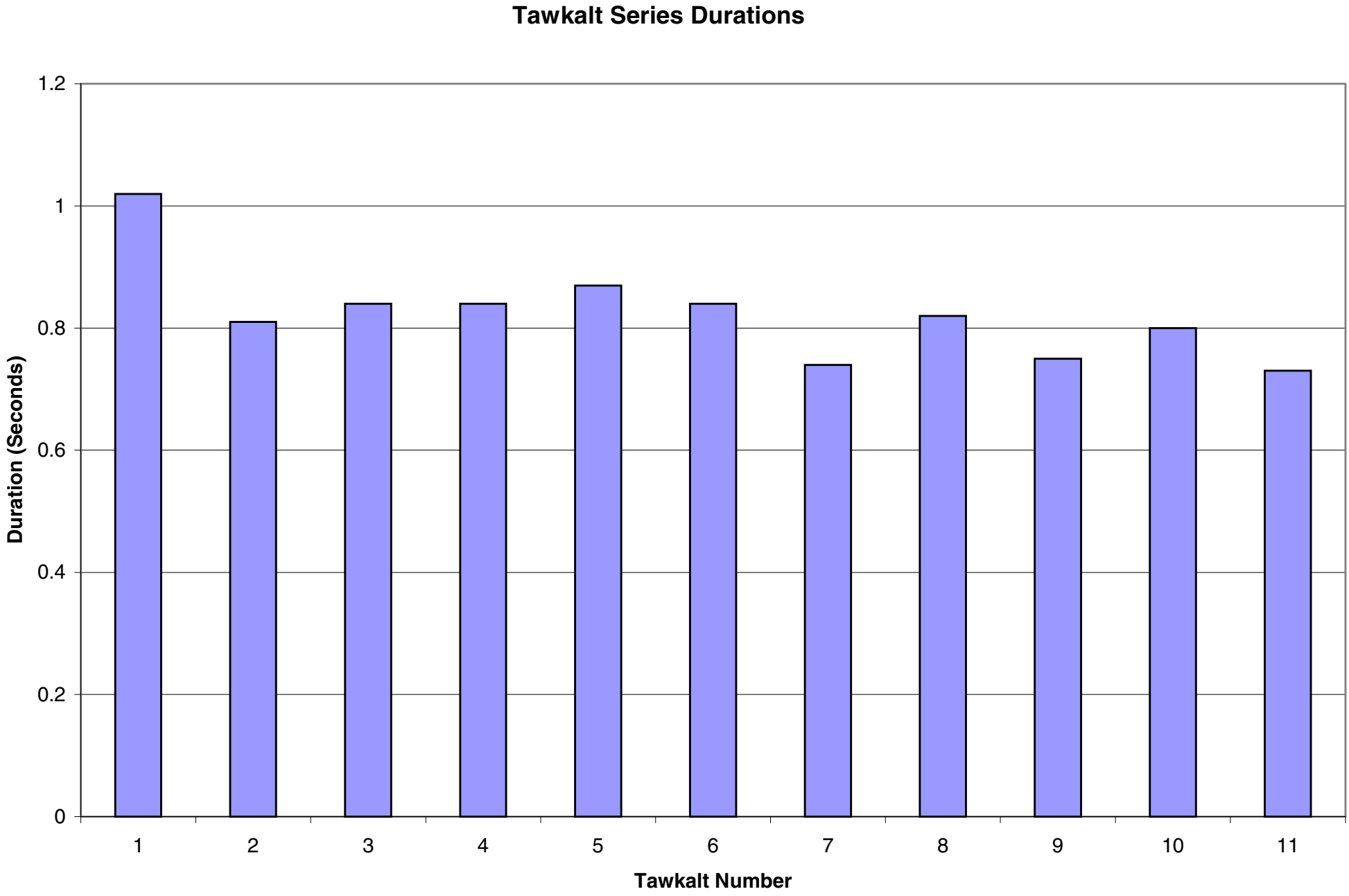
Tawkalt Series Durations
1.2
1
0.8
0.6
0.4
0.2
0
Duration (Seconds)
1
2
3
4
5
6
7
8
9
10
11
Tawkalt Number

FIGURE 4

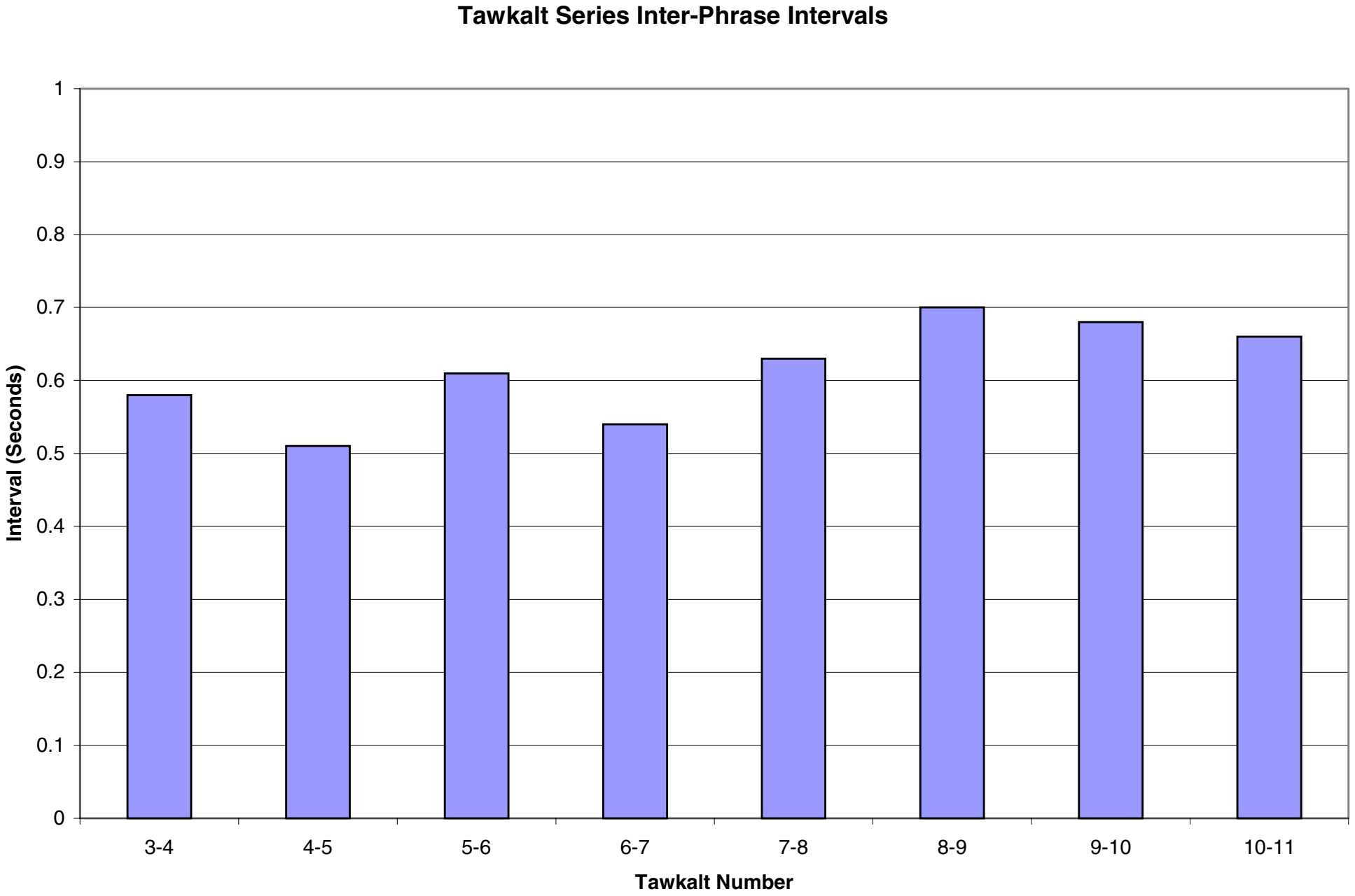
Tawkalt Series Inter-Phrase Intervals
Interval (Seconds)
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
3-4
4-5
5-6
6-7
7-8
8-9
9-10
10-11
Tawkalt Number

FIGURE 5