

[54] **APPARATUS AND METHOD FOR RETARDATION OF RECORDED SPEECH**
 [76] Inventor: **David Thurston Griggs**, 5128 S. Rolling Rd., Baltimore, Md. 21227
 [22] Filed: **May 12, 1972**
 [21] Appl. No.: **252,568**
 [52] U.S. Cl. **179/1 SA**
 [51] Int. Cl. **G101 1/16**
 [58] Field of Search 179/1 SA, 15.55 R, 15.55 T, 179/100.2 T

2,115,803 5/1938 Dudley 179/15.55 R
 2,286,072 6/1942 Dudley 179/15.55 R
 3,584,158 6/1971 Jefferies 179/1 SA
 2,903,521 9/1959 Ellison 179/15.55 R
 3,555,203 1/1971 Scott 179/15.55 R

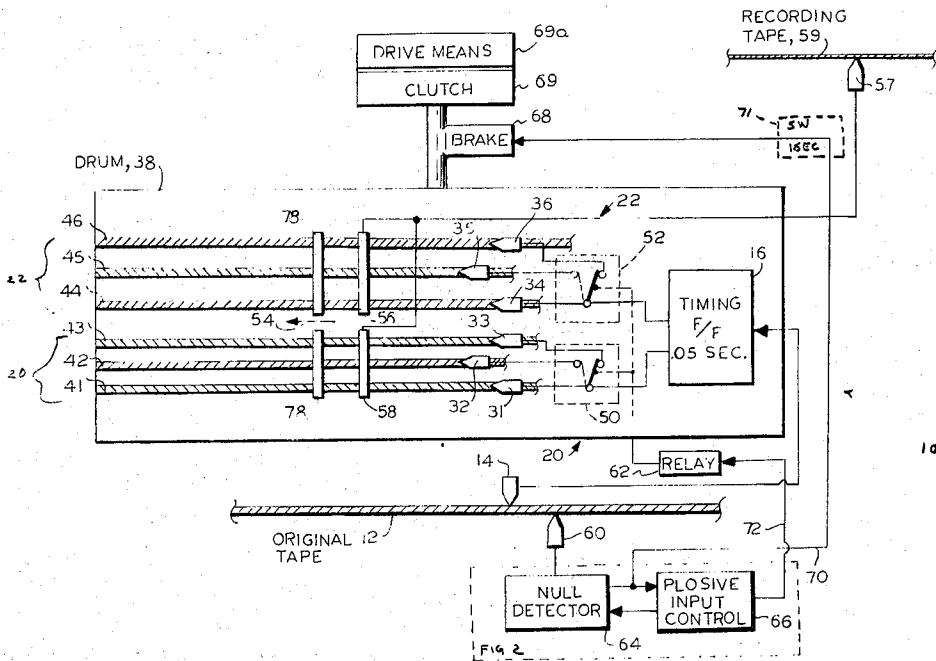
Primary Examiner—Kathleen H. Claffy
 Assistant Examiner—Jon Bradford Leahey
 Attorney, Agent, or Firm—Misegades, Douglas & Levy

[57] **ABSTRACT**

A method and system for translating or processing recorded speech to a slower rate of speech while retaining its original frequency values, with interposition of compensations which include intensification of "plosive" and other sounds which follow silences or breaks in the stream of speech.

3 Claims, 2 Drawing Figures

[56] **References Cited**
UNITED STATES PATENTS
 3,621,150 11/1971 Pappas 179/1 SA
 3,532,821 10/1970 Nakata 179/1 SA



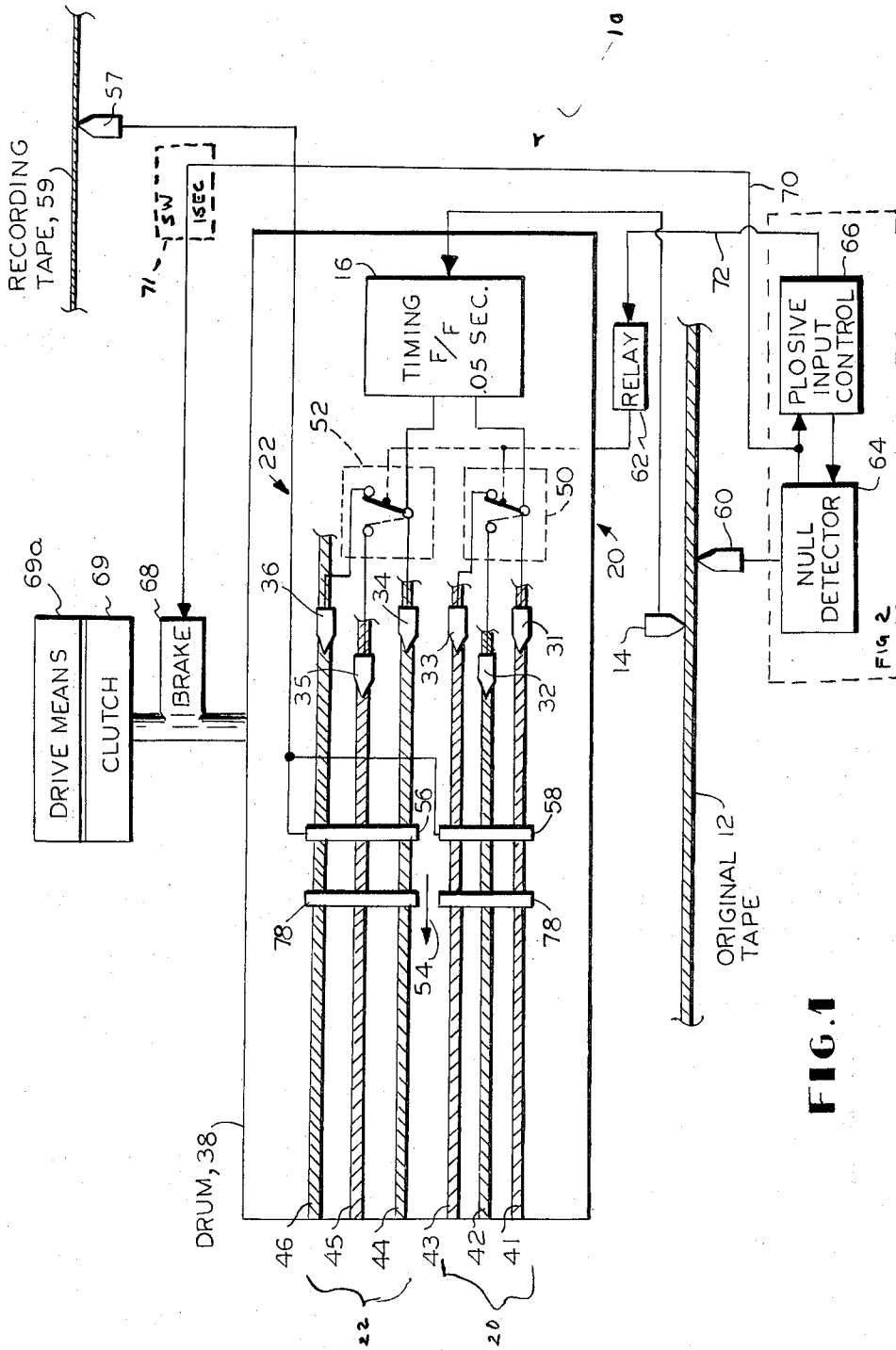


FIG. 1

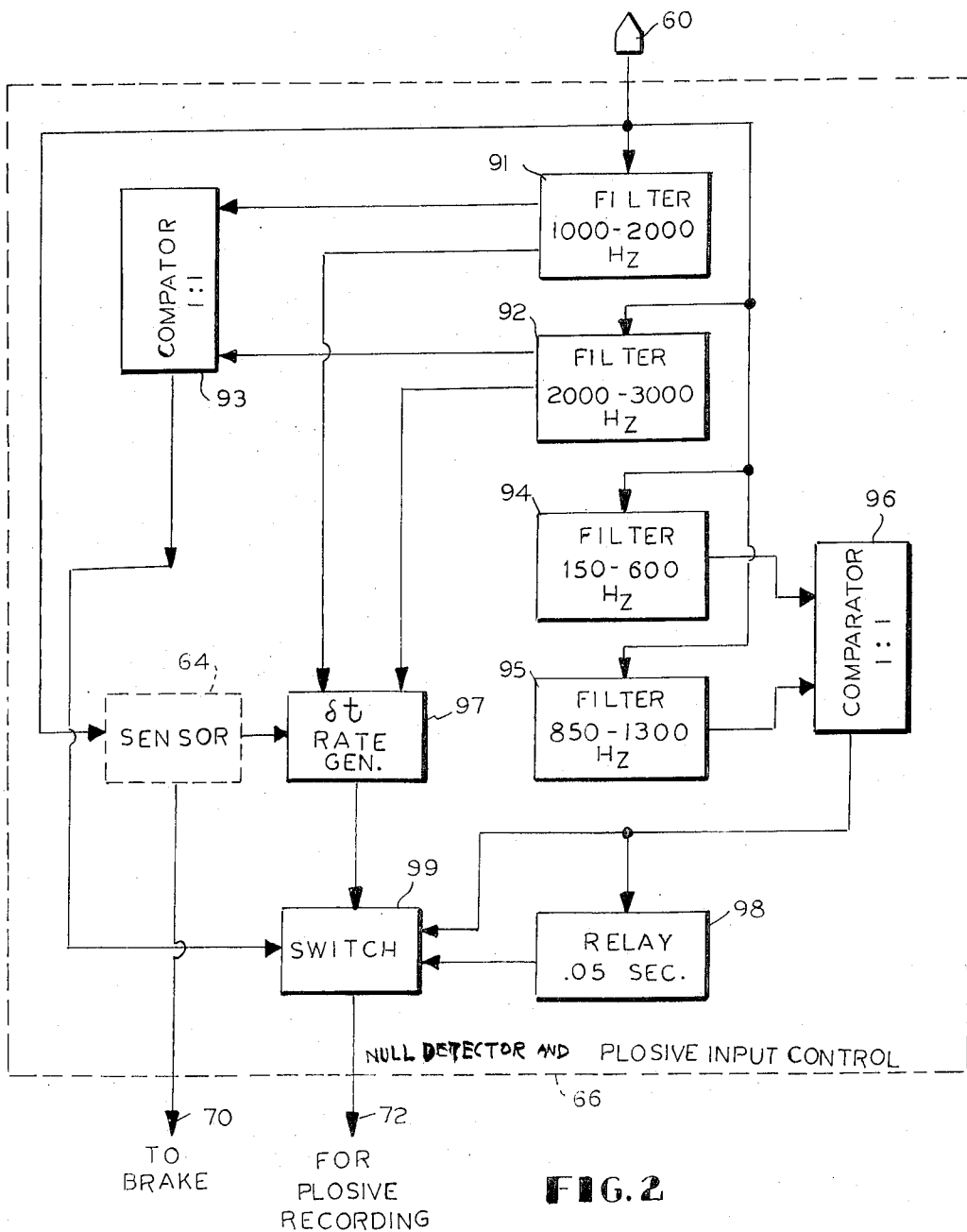


FIG. 2

APPARATUS AND METHOD FOR RETARDATION OF RECORDED SPEECH

CROSS-REFERENCES TO RELATED APPLICATIONS

The present invention is a related invention of the inventor's applications:

Serial No.	Title	Appl. Filing date
1,739 (now U.S. Pat. No. 3,646,576)	Speech Controlled Phonetic Typewriter Sound Separator for Talkwriter	Jan. 9, 1970
86,868	Real-Time Mechanical Detection of Diphthongs for Talkwriter Apparatus	Nov. 2, 1970
177,764	Mechanical-Stenographic Apparatus for Transcribing Group Proceedings	Aug. 13, 1971
252,569		May 12, 1972

BRIEF SUMMARY OF THE INVENTION

The present invention relates to methods and means for multiplying or extending the duration interval of recorded speech through re-recording, to the end that there is retained the essence of the original frequencies and time-values for subsequent mechanical analysis or processing, and for other general purposes.

More particularly, the invention relates to means for intensifying plosive input sounds in a re-recording process of decelerating the rate of speech and interposing efforts to preserve the stop or silence signal, actually an absence of signal in the usual sense, and to amplify the effective plosive input that follows it by intensification through electronic means. The plosive sound identifies the release of the stop, and the electronic means provides for amplification of this effect and introducing it at the appropriate time onto the recorded speech or voice.

Another particular feature of the invention is directed to means for re-recording by eliminating non-useful or irrelevant intervals of silence not connected with plosive speech sounds.

It is one of the features and advantages of the invention to provide a system of prolongation or delay of the components of a voice signal to the end that there is maintenance and preservation of the quantitative and qualitative components of the entire speech spectrum as recorded. The voice signal is prolonged to the end that there is no breaking down of the speech signal until it is subsequently introduced into a talkwriter.

FIELD OF THE INVENTION

When speech comes too fast, or when a slow-motion version of it is needed, such as for mechanical processing or analysis, there is presently no known means for slowing it down or attenuating its phonetic elements for analysis other than by spectrograph or by distortion of the frequency and time-value. Interpolation of brief bursts of silence has been used for attenuation when human listeners are involved, but that process will not actually continuously prolong the sonic material, and therefore will not significantly aid mechanical analysis of the speech for introducing directly into a talkwriter. In many applications, the spectrograph is too costly and cumbersome and takes too long in providing an analyti-

cal breakdown of fast speech. In other instances, it is necessary to preserve the original frequencies and time values rather than to distort them by modifying the playback speeds. The invention is intended to attenuate the phonetic characteristics at the original frequencies, and to permit conversion of the original speech track to a faithful sonic analog that takes about 75 - 100 percent longer. Although the actual effect is that of slowing down the material to half-time, still because interstitial silences are deleted in the proposed process, it is found that less than double time may be required.

The present invention divides the speech input, such as from a tape, into intervals of 1/20th of a second and repeats each one twice in succession, making it by synthesis into twice the duration of the original. It is not necessarily separate speech phonemes themselves that are thus treated, since they cannot be readily separated on a tape, but the process multiplies whatever sonic material occurs in each successive 0.05 second. An exception is the case of plosive sounds. Since it is not possible to repeat a plosive release to any advantage in making machine analysis, and since consecutive indications of a plosive would be undesirable for some applications, as to a talkwriter for example, the invention proposes that the plosive releases be recorded at double strength rather than be repeated. Provision is made to delete periods of silence from the conversion on the assumption that machine transcriptions will follow, capable of handling continuous inputs.

By addition of more of the same kinds of components, the device of the invention can be enlarged to give three times the attenuation of speech rather than twice. This would require, for the intermediate recording process, four tracks for each approximately 0.05-second segment rather than three with a change of rotational speed of the drum and/or size of the recording heads. The preferred embodiment of the invention seeks to exemplify these features and realize these objects.

BRIEF DESCRIPTION OF THE FIGURES OF THE DRAWINGS

The above and other objects and advantages of the invention will become apparent upon full consideration of the following detailed description and accompanying drawings in which:

FIG. 1 is a typical schematic diagram of the system of the invention for achieving retardation of recorded speech within the purview of the invention; and

FIG. 2 is a circuit or block diagram of the plosive input control shown as a component in FIG. 1.

DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT

Referring now to the drawings, there is shown a sound rate retardation system 10 having a main data or voice storage tape 12 (disposed on a drum 38 as described below) carrying or storing a fast-speech signal thereof as it moves right to left past a detector or pick-off head 14 for conduction to a cycling or timing switch 16 that switches the fast-speech signal to one of two sets 20,22 of three recording heads 31,32,33; 34,35,36 for being recorded on associated tapes 41,42,43; 44,45,46 on a drum 38. The heads are connected through a set of ganged switches 50,52.

The tapes 41-46 move at a relative constant velocity of, for example, 7½ inches per second in a direction

shown by arrow 54. The sets 20,22 may each comprise a single wide tape having the three magnetic recording tracks of each set, or the tapes may comprise six actual tapes.

The recording heads 31,33 and 34,36 of the sets 20,22 are electrically acoustically symmetrically aligned for simultaneous recording, and the other heads 32,35 are aligned so that each will record also simultaneously but at an advanced position along the tape set. The heads 32,35 are so positioned as shown to provide recording of signals in time-advanced relation along the tape considering the direction of the tape set movements. For playback purposes (not shown in detail) the two paired heads 34,36 will read at a time interval on the tape of, for example, 0.05 seconds later and immediately following the signal recorded by the single head 35; and the paired heads 31,33 will read back at a time interval on the tape of 0.05 seconds later and immediately following the signal recorded by the single head 32.

The input from the pick-off head 14 is switched in 0.05 second intervals by the cycling or timing switch 16, which may be a flip-flop, or similar type electronic type switch, alternatingly to the ganged switches 50,52. Thus by the positioning of the heads 32,35, each set provides a total of about 0.1 seconds of newly recorded linear tape time utilized to record alternative 0.05 second segments of the original main data tape.

When there is a fast-speech input from the main data or voice storage tape 12, it will be recorded in two successive 0.05 time segments by heads 31 and 32 or 34 and 35, and no other sonic data or material is recorded or impressed on the tape sets 20,22. The successively recorded sound is synthesized by pickup heads 56,58 converting the sequential signal storage data of tapes 41 and 42 or 44 and 45 into slow speech sound which may then be recorded by head 57 on recording tape 59.

When there is a short interval of no sound or of a stop for plosives, this initial condition is detected in advance by use of a pick-off head 60 that serves as a detector of silence by null detector 64 that produces a no-signal or gate output indicative of such silence, which gate output activates a plosive input control to activate a relay 62 for activation of the gang switches 50,52 to the right position for accomplishing the effect of recording by the aligned heads 31,33 or 34,36 of each of the sets; this leaves a silence in the preceding 0.05-second period of playback on set 20 or 22 and a dual recording by either of these paired heads 31,33 or 34,36 of what is in the 0.05-second period of plosive release, doing so on two separate tapes of the respective set. This requires that the pick-offs be spaced as shown in FIG. 1 where distances shown are proportional and approximately actual size for operation at 7½ inches per second. The sets of pick-ups 56 and 58 are interconnected and fed to a recording head 57 for a new tape 59 onto which the attenuated and synthesized speech is recorded or transcribed.

The position of the head 60 is 0.05 seconds ahead of the pick-off head 14.

The advance pick-off head 60 feeds a silence signal to a gate or null detector 64 to which output is connected the plosive input control 66, which is a detector for a rate of change of input sound-to-sound signal. The null detector 64 on periods of silence provides a signal on line 70 to activate the brake 68 for instantaneously stopping the tapes 41-46 on drum 38 from being driven

through clutch 69 by drive means 69a as shown, which continues until release of the non-input condition in null detector 64. On line 70 conveying a non-input signal, a time delay switch 71 of about 1 second's duration blocks application of the brake 68 unless the period of silence exceeds one continuous second of time, in each instance.

At the release of the non-input condition, plosive input control 66 provides an output signal on line 72 to relay 62 to activate the gang switches 50,52 to the right position to activate simultaneous dual recording of the first sounds that then and there transpire following a period of silence, which is when the rate of change sensor circuit in the plosive input control (FIG. 2) shows a sudden gain or the presence of a plosive.

In the case of unvoiced plosive speech sounds, these will be the releases of breath in those sounds by which they can be identified, and a dual recording of them will have been made. Similar amplification of the initial sounds following any silence also can prove to be useful if there is a possibility of their being confused with plosives. With voiced plosives, although there may be no stoppage of the drum 38 while voicing is active (and is being attenuated by this instrument), the release will be intensified by the sensor of plosive input control 66 and its control over switches 50 and 52.

The drive means 69a for the tape must be clutch-mounted by clutch 69 so that instantaneous stopping and starting in connection with the brake will be possible.

After the tracks of tapes 41-46 have been recorded using the multiple heads 31-36, they can be played off subsequently at unison speed by pick-ups 56 and 58 aligned uniformly, so that each reads the proper consecutive period of 0.05 seconds in the order in which they were originally recorded.

Sets of erasing heads 78 can follow each of the sets of pick-offs 56 and 58, if desired, to prepare the tape tracks on the drum 38 for subsequent re-use.

Details regarding the null detector and plosive input control 66 of FIG. 2 are as follows: the pick-off 60 that is positioned 0.05 seconds advance of pick-off 14 is supplied to a network of four filters 91,92,94,95, and by subsequent processing gives two outputs: (a) a non-input signal 70 to operate the brake 68, and (b) a switching signal 72 for plosive recordings as described above. The output 70 comes simply from a silence-detector or gate 64 which leaves a circuit closed except when excited. The plosive control switch 99 operates only when satisfied by either one or two sets of dual requirements: (a) there must be a rate of change that has been preset to show the rapid bursts of a plosive in all cases, and (b) this must be accompanied by either an indication that the total amplitude of the bandwidth 1,000 to 2,000 Hz of filter 92 is not greater than that of bandwidth 2,000 to 3,000 Hz of filter 92 at the time the burst starts, or that the total amplitude of bandwidth 150 to 600 Hz of filter 94 is greater than that of bandwidth 850 to 1,300 Hz of filter 95. The first situation satisfies the condition of a plosive burst rather than a vowel suddenly released following silence; The second situation satisfies the condition of a voiced plosive by indicating concentrated glottal resonances which accompany a voiced plosive. Since that intensified resonance occurs either before or during the release or burst, provision is made by means of a 0.05 second relay 98 for the indication to be supplied again in the

following period of 0.05 second of analysis when the requisite rate of change of generator 97 takes place.

The outputs of filters 91 and 92 are compared by comparator 93 to ascertain whether the total amplitude of bandwidth 1,000-1,999 Hz is not greater than that of 2,000-3,000. If it were greater, then a vowel or nasal might be present rather than a plosive release. The outputs of filters 94 and 95 are compared in comparator 96 to see if the total amplitude of the lower frequencies is greater. If so, this will mark glottal intensification in circumstances peculiar to voiced plosives, and the switch 99 can be activated accordingly, if there is an appropriate rate of change as well. The selection of frequency bandwidths in this instance has been made so as to exclude fricative, nasal and vowel characteristics.

Additional embodiments of the invention in this specification will occur to others and therefore it is intended that the scope of the invention be limited only by the appended claims and not by the embodiments hereinabove. Accordingly, reference should be made to the following claims in determining the scope of the invention.

What is claimed is:

1. A data processing method for converting voice signals to speed-retarded voice signals for alpha-numeric print-out and for other means, comprising the steps of:

1. sensing a voice signal,
2. cycling the signal in essentially 0.05 second segments for two sets of recorder heads in 0.05 second displacement along a recording tape,
3. detecting a null input in essentially 0.05 second advance of the null reaching the sensing step and producing a gate output when no signal is detected,
4. sensing a rate of change in detecting a null input to activate a brake to stop drive means for said recording tape until change of the non-input condi-

tion.

5. driving gang switch means upon change of the non-input condition to provide simultaneously dual recording for a period of the first sounds that transpire following a period of silence.

2. Data processing means for converting voice signals to speed-retarded voice signals for alpha-numeric print-out and for other means comprising: voice signal sensing means:

a signal cycling means for receiving the output of the voice signal sensing means and dividing the voice signal into time divided components of essentially equal duration, recording means to store the divided components on tracks for each series of the divided components in time-sequential segments of storage means;

a simultaneous sound pick-off means for taking each divided component from the storage means for synthesizing into a speech-retarded voice signal;

a rate of change in voice to null input means sensing in a time interval in advance of said voice signal sensing means that a null input is approaching the voice signal sensing means; and

a driving gang switch means responsive to a non-input condition providing simultaneously dual recording in the storage means for a period of the first sounds that transpire following a period of silence and inactivate the time-sequential recording function of the recording means, so that the simultaneous sound pick-off means takes each divided component from the storage means for synthesizing into an emphasized phoneme voice signal.

3. The invention of claim 2 wherein the time-durations of the cycling means and time-sequential function are of a magnitude of essentially 0.05 seconds.

* * * * *

40

45

50

55

60

65