

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
10 September 2004 (10.09.2004)

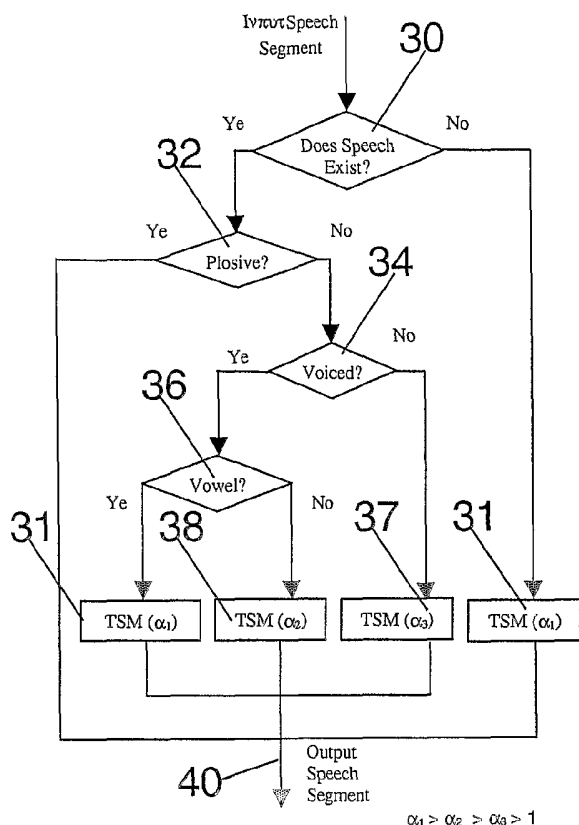
PCT

(10) International Publication Number  
WO 2004/077381 A1

- (51) International Patent Classification<sup>7</sup>: **G09B 19/04**, G10L 21/04, 13/02
- (21) International Application Number: PCT/IE2004/000029
- (22) International Filing Date: 27 February 2004 (27.02.2004)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 0304630.7 28 February 2003 (28.02.2003) GB
- (71) Applicant (for all designated States except US): **DUBLIN INSTITUTE OF TECHNOLOGY**; Fitzwilliam House, 30 Upper Pembroke Street, Dublin 2 (IE).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): **DONNELLAN, Olivia** [IE/IE]; Dunmore Road, Tuam Road, Co. Galway (IE). **COYLE, Eugene** [IE/IE]; 64 Bothar Brugh, Drogheda, Co. Lough (IE). **JUNG, Elmar** [DE/IE]; Grove House, Hianault, Foxrock, Dublin 18 (IE).
- (74) Agent: **HANNA, MOORE & CURLEY**; 11 Mespil Road, Dublin 4 (IE).
- (81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

[Continued on next page]

(54) Title: A VOICE PLAYBACK SYSTEM



$\alpha_1 > \alpha_2 > \alpha_3 > 1$

(57) Abstract: The present invention relates to the fields of voice play back and digital signal processing and in particular applied to the field of language learning. Computer Assisted Language Learning (CALL) packages are software packages used by students to assist them in learning a language. A significant disadvantage of existing CALL packages is that the speech clips used tend to be clips recorded using speakers who are speaking slowly and deliberately. The CALL package merely plays the selected clips back to the user. The resulting speech tends to be somewhat artificial, i.e. the speech is deliberate. The resulting speech is also clearer and much more defined than the conversational, idiomatic speech one would expect to hear on the streets or in everyday life. This results in the problem that students using only these packages have a low level of recognition of colloquial, "everyday" speech. The present invention overcomes this disadvantage by providing a voice playback system having an input means for accepting a segmented speech clip and a characteristic associated with each segment of the speech clip. A time scale modification means alters the duration of individual segments of the speech clip, wherein the level of alteration for the individual segment is determined with reference to the associated characteristic of the individual segment. The altered clips are then combined into an altered duration speech clip.

WO 2004/077381 A1



(84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

**Published:**

— with international search report

## A VOICE PLAYBACK SYSTEM

### Field of the Invention

5           The present invention relates to the fields of altered speed voice play back and digital signal processing as might be applied to the field of language learning.

### Background

10           Computer Assisted Language Learning (CALL) packages are software packages used by students to assist them in learning a language. A CALL system, as known from the prior art and illustrated in simplified form in FIG. 1, comprises a CALL software package 1, which may be run on a conventional computer, for example a personal computer using the MICROSOFT WINDOWS operating system.

15           The CALL package has an associated datastore of speech clips 2 which may be downloaded to a user's PC from the Internet and/or stored on a storage medium within the computer, for example on the hard disk, or on a removable storage medium, for example a CD or a DVD.

            Typically, the CALL packages allow users to interact with the system, by  
20           providing an input 3 through a keyboard or mouse. Thus, for example, users may select a particular speech clip, which may be identified as a lesson, using a menu in the CALL package. The CALL package in turn retrieves the selected speech clip from the datastore and a voice playback system 4 plays the retrieved speech clip through a loudspeaker or other audio output device. In addition to playing back speech clips,  
25           CALL packages may also provide other features, for example an associated video clip or text information to assist the user. Additional information could include for example text representing the content of the speech clip.

            One of the major aims of CALL packages is to enhance the students  
            perception, and hence comprehension, of the foreign language. In the acquisition of a  
30           second language, the ability to slow down natural speech can prove very beneficial. Research has indicated that a learner's ability to process spoken language is affected by the rate of delivery. It is believed that slow colloquial speech is the only practical model to be used in listening exercises. However, a significant disadvantage of

existing CALL packages is that the speech clips used tend to be clips recorded using speakers who are speaking slowly and deliberately. The CALL package merely plays the selected clips back to the user. The resulting speech tends to be somewhat artificial, i.e. the speech is deliberate. The resulting speech is also clearer and much more defined than the conversational, idiomatic speech one would expect to hear on the streets or in everyday life. This results in the problem that students using only these packages have a low level of recognition of colloquial, 'everyday' speech.

In particular, it would be of considerable benefit if a system could be provided which would allow a user to speed up and/or slow down a piece of recorded speech when listening to recorded colloquial, 'everyday' speech without losing significant quality or content.

US 6,070,135 discloses a method using pitch detection for distinguishing non-sounds from voiceless sounds. This method may be used in the reproduction of speech signals at varied play-back speed.

US2002101368 discloses a method for altering the speed of audio in an MPEG decoding system in accordance with a user's designated playback speed for video content. In particular, a TSM algorithm is performed on the audio data of the input queue to decrease the quantity of the audio data when the designated playback speed is faster than a normal playback speed or to increase it when the designated playback speed is slower than the normal playback speed, in accordance with a value of the designated playback speed. The effect of which is that tone of the altered speed audio is substantially identical to that of the normal playback speed and thus is less obtrusive to the listener than altered tone content. Whilst the quality of the altered speed audio may be suitable for use in such situations where a user is fast forwarding a video segment, it is not of sufficient quality to be particularly suited to language learning applications. US5809454 discloses a method to solve the same problem addressed by US2002101368 for altering the speed of audio in an MPEG decoding system in accordance with a user's designated playback speed for video content. One method disclosed increases\decreases the duration of soundless intervals to effectively speed up or slow down the speed of playback.

WO02082428 discloses a technique utilising Time Scale Modification (TSM) which is primarily intended for use with coding and decoding techniques. In particular, it describes the use of one TSM algorithm, referred to as the synchronised

overlap-add (SOLA), which is an example of a waveform approach algorithm. It describes a particular problem with SOLA when applied to speech in that SOLA introduces unwanted artefacts into speech when decoded. In particular, the method analyses individual frames of signals for voiced and un-voiced components with  
5 different expansion or compression techniques being utilised for the different types of signal. Whilst this may be adequate for general speech coding, its quality of result is not adequate for use in fields such as language learning where listeners are paying particular attention to the speech.

EP0817168 describes a similar approach in which an input audio signal is  
10 transmitted from a memory device to a voiced sound/unvoiced sound deciding portion. In the voiced sound/unvoiced sound deciding portion, a decision is made as to whether the input sound signal is a voiced sound or an unvoiced sound. A speech velocity converter outputs the unvoiced sound as it is and a time compression is carried out so as to output the voiced sound. Again, this method suffers from  
15 substantially the same disadvantages as described for WO02082428.

US5828994 discloses a method for non-uniform time scale modification of recorded audio. The method alters the compression rate so that greater compression is applied to the portions of the speech which are least emphasised by the speaker and less compression is applied to the portions of the speech which are most emphasised.  
20 The patent discloses that the calculation of the emphasis may be done using a speech to text converter in combination with an emphasis dictionary or more practically using the energy content of the signal. The present invention seeks to improve upon the teaching in this document.

25

### **Summary of the Invention**

Accordingly, the present invention seeks to provide an improved CALL system and in particular one which facilitates students learning colloquial speech. Thus, in one aspect of the present invention, a voice playback system is provided  
30 comprising an input means for accepting a segmented speech clip and a characteristic associated with each segment of the speech clip, time scale modification means for altering the duration of individual segments of the speech clip, wherein the level of alteration for the individual segment is determined with reference to the associated

characteristic of the individual segment, and an output means for combining the altered individual segments into an altered duration speech clip. The time scale modification means uses a time-domain overlap-add technique to alter the duration of individual segments. Suitably, the characteristic is selected from one or more of the following: a plosive, vowel, voiced consonant, unvoiced consonant or silence. In the case where the associated characteristic is a plosive the duration of the segment may be left unchanged. The time-domain overlap-add technique may be a synchronised overlap-add algorithm. Alternatively, the time-domain overlap-add technique may be an adaptive overlap-add algorithm. The level of alteration to be applied to individual segments may be determined with reference to a user input associated with a desired playback speed.

Preferably, the level of alteration where the characteristic is a vowel is greater than the level of alteration where the characteristic is that of a voiced consonant.

Suitably, the level of alteration where the characteristic is a voiced consonant is greater than the level of alteration where the characteristic is an unvoiced consonant. The level of alteration applied where the characteristic is a vowel may be substantially equal to that applied where the characteristic is a silence. Suitably, segments having a vowel characteristic are time scale modified at a rate of  $\alpha_1$ , segments excluding plosives having a voiced consonant characteristic are time scale modified at a rate of  $\alpha_2$ , and segments excluding plosives having a unvoiced consonant characteristic are time scale modified at a rate of  $\alpha_3$ , where  $\alpha_1 > \alpha_2 > \alpha_3$ . Moreover,  $\alpha_3$  may be greater than 1.

In one aspect of the invention, there is a substantially linear relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ , whilst in an alternative there is a substantially exponential relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .

Desirably, the voice playback system described may be applied within a computer assisted language learning system.

The invention also extends to a computer readable storage medium having stored thereon at least one segmented speech clip comprising a plurality of segments, each segment having a characteristic associated with it, wherein the at least one characteristic identifies the suitability of the segment for time scale modification.

In another embodiment, the invention provides for a method of altering the duration of

a segmented speech clip comprising the steps of: altering the duration of individual segments of the speech clip using time scale modification, wherein the level of alteration for the individual segment is determined with reference to an associated characteristic of the individual segment, and combining the altered individual segments into an altered duration speech clip. In this method, a time-domain overlap-add technique may be used for time scale modification. Examples of a time-domain overlap add technique include synchronised overlap-add algorithm and adaptive overlap-add algorithm techniques. The level of alteration to be applied to individual segments may be determined with reference to a user input associated with a desired playback speed. The characteristic may be selected from one or more of the following: a plosive, vowel, voiced consonant, unvoiced consonant or silence. In situations where the associated characteristic is a plosive, the duration of the segment is preferably left unchanged. Suitably, the level of alteration where the characteristic is vowel is greater than the level of alteration where the characteristic is that of a voiced consonant. Desirably, the level of alteration where the characteristic is a voiced consonant is greater than the level of alteration where the characteristic is an unvoiced consonant.

Desirably, the level of alteration applied where the characteristic is a vowel is substantially equal to that applied where the characteristic is a silence. In a preferred method, segments having a vowel characteristic are time scale modified at a rate of  $\alpha_1$ , segments having a voiced consonant characteristic are time scale modified at a rate of  $\alpha_2$ , and segments having a unvoiced consonant characteristic are time scale modified at a rate of  $\alpha_3$ , where  $\alpha_1 > \alpha_2 > \alpha_3$ .  $\alpha_3$  may be greater than 1. In one configuration, there is a substantially linear relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ . In an alternate configuration, there is a substantially exponential relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .

The method is particularly beneficial for use in a computer assisted language learning environment. In this regard, the method of the present invention may be used to produce altered duration speech, which may be recorded on a storage medium, tape, CD, DVD, etc for subsequent playback by students.

The method described may be embodied as code within a computer program product which, when run on a computer, effects the running of the method.

The computer program described may further comprise at least one segmented speech clip comprising a plurality of segments, each segment having a characteristic associated with it, wherein the at least one characteristic identifies the suitability of the segment for time scale modification.

- 5 The method of altering the duration of a speech clip extends to the additional steps of: segmenting the speech clip, associating a speech characteristic to each segment to produce a segmented speech clip and applying the methods described above to produce an altered duration speech clip.

A computer readable storage medium having stored thereon at least one segmented  
10 speech clip comprising a plurality of segments, each segment having a characteristic associated with it, wherein the at least one characteristic identifies the suitability of the segment for time scale modification. The characteristic may be selected from one or more of the following: a plosive, vowel, voiced consonant, unvoiced consonant or silence.

15

Other aspects and advantages of the invention will be appreciated from the detailed description which follows.

#### **Brief Description of the drawings**

20 The invention will now be described in greater detail with reference to the accompanying drawings in which;

FIG. 1 is an exemplary prior art CALL system,

FIG. 2 is CALL system according to the present invention,

FIG. 3 details an algorithm suitable for use with the present invention,

25 FIG. 4 is a detailed diagram of an exemplary voice playback system suitable for use in the CALL system of FIG. 2,

FIG. 5 is a diagram of a datastore suitable for use with the CALL system of FIG. 2,

FIG. 6 is a flowchart of a method suitable for use in the voice playback system  
30 of FIG. 4,

FIG. 7 is a graph providing a comparison of results vis a vis the prior art and the present invention, and



FIG. 8 is a graph providing a further comparison of results vis a vis the prior art and the present invention.

### Detailed Description of the Invention

5

The inventors of the present invention have realised that the key to an effective CALL package is the use of slowed speech, rather than a slow speaker. Moreover, they believe that, by slowing down natural, 'everyday' speech, a better training source is provided than that of using a slow speaker.

10

This is achieved, as shown in FIG. 2, using a database/datastore 12 of natural speech clips and a process of time scale modification (TSM) when the voice playback system 11 within the CALL package 10 is playing back a clip. TSM refers to the process of altering the duration of an audio segment. A signal may be expanded, producing a signal of longer duration (a slower signal), or compressed, resulting in a signal of shorter duration (a faster signal). To be properly time-scaled, the modified signal should maximise the retention of all of the characteristics of the original signal. For example, the perceived pitch and naturalness should be maintained. Simply adjusting the playback rate of the signal will alter the duration of the signal, but will also undesirably affect the frequency (pitch) content.

15

20

Of the current methods available for performing TSM, many are capable of producing an output which may be applied in a basic CALL package 10. However, it is preferable for effective CALL packages, that the quality is extremely high, with any distortion or unnaturalness minimised. Otherwise, the value of the CALL package 10 is limited.

25

Examples of time-scaling techniques which may be applied to the present invention include time-domain overlap-add techniques (TDOLA), frequency-domain techniques, and parametric techniques. However, the TDOLA approach is particularly suited to periodic signals such as voiced speech, and provides a good compromise between quality and efficiency, i.e. providing a high quality output for a relatively low computational load.

30

A TDOLA technique performs time scale modification essentially by duplicating small sections of the original signal, and adding these duplicated segments using a weighting function, i.e., to make a signal have a longer duration, individual

small segments are made longer by duplicating or repeating them. The technique requires firstly segmenting the waveform into a series of overlapping frames by windowing the signal at intervals along the waveform. These frames can then be added together, but with a different amount of overlap. Time-scale expansion is  
5 achieved by creating a waveform through the recombination of frames with a reduced amount of overlap. The amount of overlap required depends upon the desired expansion. Similarly, an increased amount of overlap will result in a time-scale compressed signal. The different TDOLA techniques generally vary in the way the waveform is segmented (choice of window, where segmentation occurs etc), or how  
10 successive frames are overlapped or aligned.

A commercially popular TDOLA algorithm is the Synchronised Overlap-Add (SOLA) algorithm (as described in Roucus, S. and Wilgus, A.M., "High-Quality Time-Scale Modification for Speech", IEEE Proceedings on Acoustics, Speech and Signal Processing, March 1985, pp. 493-496), because of its low computational  
15 burden with relatively high quality output. A more recent development (Lawlor, Bob, Audio Time-Scale and Frequency-Scale Modification, PhD Thesis, Department of Electrical and Electronic Engineering, University College, Dublin, November 1999), is the Adaptive Overlap-Add (AOLA), which offers an order of magnitude saving in computational burden without compromising the output quality, making it particularly  
20 suitable for real-time implementation in applications such as a CALL package 10. The inventors of the present invention are grateful for the assistance of the above referenced author (Bob Lawlor) in the initial stages of the research project, which has led to the development of the present invention.

25 The AOLA algorithm will now be explained, with reference to FIG. 3, in the following manner:

1. A window length of  $\omega$  is chosen such that the lowest frequency component of the signal will have at least two cycles within each window (FIG 3a),
2. The frame is duplicated and the duplicate of the original is shifted to the right  
30 to align the peaks (FIG 3b),
3. Overlap-adding the original frame and its duplicate produces a naturally expanded waveform; (FIG 3c), the length of this expanded segment is  $\omega \cdot ne$ , where  $ne$  is the natural expansion factor.

4. A portion of length *step* of the input signal is taken and is concatenated with the last expanded segment; (FIG 3d-e), *step* varies for each iteration and is a function (1) of  $\omega$ ,  $\alpha_{ne}$  and  $\alpha_{de}$  desired expansion factor

$$step = \omega \frac{(1 - \alpha_{ne})}{(1 - \alpha_{de})} \quad (1)$$

5. The next segment to be analysed is the length frame ending at the right edge of the appended segment, (FIG 3e). This process continues until the end of the input signal is reached.

This method has a low computational load relative to other commercial algorithms of similar quality. Another advantage of the method is that there are no discontinuities at the frame boundaries, as can be the case in other algorithms. This is because, referring again to (FIG. 3), the area in (c) ending in the vertical dashed line and the area ending in the vertical dashed line in (d) are exactly the same shape, so the segment appended to the expanded waveform will be aligned perfectly (e).

Although TDOLA techniques provide the best compromise of computational load and quality, they have a few problems. As previously mentioned, TDOLA techniques perform time scale expansion by duplicating or repeating small segments of the original signal. If one of these segments to be repeated consists of a transient, such as a plosive, this may result in unwanted clicks, or what could be perceived as a 'stuttering' effect. Current TDOLA techniques assume that all segments of the speech, whether voiced, unvoiced, vowel or consonant, should be time-scaled at a uniform rate. A problem with this is that the transients (e.g. plosives in speech or drumbeats in music) are time expanded to the same degree as non-transient segments of the original signal. If a plosive were expanded using the TDOLA technique, the sound of the plosive would be distorted, and intelligibility of the resulting speech will be diminished. Also, vowels tend to be more influenced by speaking rate than consonants. To maintain intelligibility and naturalness, different time scaling factors need to be applied to the different segments of speech.

Plosives (/b/, /d/, /g/, /k/, /p/, /t/) are produced by complete closure of the oral passage and subsequent release with a burst of air. An example of a voiced plosive is the 'duh' sound in 'dog'; an unvoiced example is 'puh' as in 'pit'. There are three distinct stages: (a) closure, when the air-stream is totally blocked by the articulators, and the air-pressure builds up behind the obstruction, (b) burst, a sudden increase in

energy when the articulators quickly open and a burst of air rushes through, and (c) transition, the transition segment to the next sound. The inventors of the present invention believe that the effect of time-scaling on plosives may be undesirable. As plosives convey a large amount of information, it is necessary to preserve their character under TSM. Moreover, they believe that at large TSM factors, plosives may be artificially transformed into fricatives, e.g. /p/ slowed down at a high TSM factor may sound more like the fricative /f/. In normal speech, the closure stage of a plosive tends to be consistent in duration, regardless of the speed of the speech. The duration of the burst also tends to be constant, and the 'suddenness' of the onset of energy needs to be maintained. Time-scaling the burst could lead to transient repetition, i.e. certain segments get repeated causing distortion and unwanted artefacts such as clicks.

The intention of the present invention is to be able to slow down speech while maintaining the naturalness and retaining all the characteristics of the original signal.

Research (Kuwabara, H., "Acoustic and Perceptual Properties of Phonemes in Continuous Speech as a Function of Speaking Rate", Proc. Eurospeech 97, pp. 1003-1006) has shown that the duration of unvoiced segments of human speech varies less than the duration of voiced segments. Other research (Ebihara, T., Ishikawa, Y., Kisuki, Y., Sakamoto, T., and Hase, T., "Speech Synthesis Software with Variable Speaking Rate and its Implementation on a 32-bit Microprocessor", 19th IEEE International Conference on Consumer Electronics (ICCE 2000), Los Angeles Airport Marriott, USA, June 2000) directed at speech synthesis suggested that a non-uniform rate be applied when time-scaling in speech synthesis so as to maintain the temporal structure of the utterance, and proposes a method of modifying only the voiced segments, or only the vowel segments in speech synthesis systems. Whilst, this information is of assistance, it does not provide a useful basis for a CALL system.

Accordingly, the inventors of the present invention have developed a voice playback system suitable for use within a CALL package which, unlike previous systems, does not treat an entire speech clip as a homogenous structure but treats the clip as a plurality of segments and determines an appropriate time scaling for each segment based on the vocal characteristic of the segment. This approach facilitates the use of natural speech clips and also provides for an improvement in quality over existing methods of altering the speed of speech playback.

The invention will now be described in greater detail with reference to FIG. 4, in which an exemplary voice playback system 11 according to the invention is shown. The voice playback system 11 has an input means/module 14 for accepting a segmented speech clip 15, i.e. an audio clip that has been segmented into a plurality of segments. Suitably, each segment in the clip contains an element of speech of a particular dominant characteristic. The characteristic may be that the element of speech is not speech (i.e. a pause between speech), plosive, voiced consonant other than voiced plosive, unvoiced consonant other than unvoiced plosive, or a vowel. This characteristic is preferably also supplied as an input 16 to the VPS. Suitably, the input module 14 is a section of software code, which is configured to retrieve the voice segments and their associated characteristics from an associated datastore 12 where audio clips and the segment characteristics are stored. An exemplary structure of a datastore 12 is shown in FIG. 5, in which an audio clip 20 has been broken up into a plurality of segments (21a, 21b,..21n) and in which each segment (21a, 21b,..21n) has an associated characteristic (22a,22b,..22n) stored against it. Whilst the segmented audio clip 20 has been illustrated as a plurality of individual segments, it will be appreciated that the individual segments (21a, 21b,..21n) of the audio clip need not be stored separately. For example, the audio clip may be stored as a single unit, with segmentation being effected using identifiers identifying the locations of individual segments within the clip (for example the start location of each segment). Each identifier and accordingly each segment may then be associated with an associated characteristic assigned identifying the characteristic of the audio content in the segment. A significant advantage of this approach of association of individual segments with a characteristic is that the association may be made at any time. For example, in the case of a computer assisted language learning (CALL) product, the characteristics may be determined by the CALL product supplier during the assembly of the CALL product. This determination of the associated characteristics may be made using a combination of automatic, semi-automatic and manual techniques readily available to the CALL product supplier. This is to be contrasted to other approaches, which alter the speed of playback using calculations or estimates made on the content at the end point of use.

It will be appreciated that suitable techniques for retrieving information from a datastore are well known in the art and that there is thus no need to describe them in detail.

Once the segments have been received by the input module 14, each of the segments is passed from the input module 14 to a time scale modification means 18 which alters the duration of the individual segments of the speech clip in accordance with algorithms known in the art as described above. It will be appreciated that the TSM means may be readily implemented in software using programming techniques familiar to those skilled in the art. However, unlike prior art systems, the present invention when performing the time scale modification on a segment, considers the characteristic 16 of the speech segment to determine the degree of TSM to be applied. The degree of TSM applied may include the non performance of TSM for particular characteristics (e.g. with plosives so as to preserve their characteristics). Similarly, higher degrees of TSM may be applied on segments containing vowels or where there is no speech present.

Moreover, the present inventors believe that as the duration of vowels are most influenced by speaking rate, so accordingly they should be time-scaled the most, say at a rate of  $\alpha_1$ . They further believe that the duration of voiced consonants varies less than vowels, but more than unvoiced consonants, therefore a rates of  $\alpha_2$  may be applied for voiced consonants, and  $\alpha_3$  for unvoiced consonants, where  $\alpha_1 > \alpha_2 > \alpha_3 > 1$  as shown in the exemplary flowchart of FIG. 6.

Thus in the method of FIG. 6, an initial determination is performed to determine whether the characteristic of an individual segment is a silence or not? If the determination concludes that it is a silence, then TSM is performed on the individual segment at a rate of  $\alpha_1$ .

In the event the determination concludes that the segment contains speech, a further decision 32 determines whether the speech content is a plosive. In the event that the speech content of the segment is a plosive, the segment is output 40 from the TSM means without any TSM. In the event that the speech content is not a plosive, a further decision 34 considers whether the speech is voiced or unvoiced. In the event that the speech content of the segment is not voiced, TSM may be applied to the segment at a rate of  $\alpha_3$  37. In the event, that the speech is voiced, a further decision

36 may be applied to determine whether the speech content of the segment is a vowel. In the event that the speech content of the segment is a vowel, TSM may be applied at a rate of  $\alpha_1$  31. In the alternative, where the segment characteristic is not a vowel, TSM may be applied at a rate of  $\alpha_2$  38. This whole process may be repeated for each segment of speech within a clip. The resulting altered duration speech segments are assembled into an overall altered duration speech clip and output by the output means 19 to an appropriate output device 5 or recording device.

A user input 3 may be included to alter the rates, thus a user may increase or decrease the speed of the speech depending on their ability to comprehend the speech. As their language skills improve users may progress to using a faster speech. The user input may be incorporated within the graphical user interface of a CALL package. For example, a slider feature could be included to identify the desired speed of playback. The rates used for a particular user input may be obtained from a look-up table or by simple scaling of basic rates using an appropriate formula.

The effectiveness of the present invention will now be described with reference to experimental results. A number of speech samples were recorded at a sampling rate of 16 kHz, and more samples were taken from the TIMIT database (DARPA TIMIT, Acoustic-Phonetic Continuous Speech Corpus. American National Institute of Standards and Technology, NTIS order number PB91-50565). An equal number of male and female speakers were used. Each signal was then analysed on a segment-by-segment basis, and manual segmental detection was performed, i.e. through a combination of examining the waveform and listening to the signal, decisions were made as to whether the characteristic of the frame consisted of a plosive, vowel, voiced consonant, unvoiced consonant or silence. The method of the present invention was then applied to produce TSM samples and compared to prior art methods as set out below in Table 1.

Method	Description
A	Uniform TSM (Prior Art)
B	Variable TSM with voiced segments only being modified
C	Variable TSM with vowels only being

	modified
D	Speech-adaptive TSM method.

Table 1

For method D, the new proposed method, two different sets of scaling parameters were considered in order to investigate the existence of a difference in quality for different sets of parameters. For each set, the requirement of  $1 < \alpha_3 < \alpha_2 < \alpha_1$  was adhered to, but the distance between the values was varied. In the first set (D1), the values varied linearly from 1 to  $\alpha_1$ . In the second set (D2), these values varied exponentially.

Each speech sample was time-scaled by each of the above methods, and at three different time-scale modification factors; 2, 2.5 and 3. Informal listening tests were used to assess the quality of the techniques. Two different forms of tests were conducted. The first consisted of 12 preference tests, in which all methods are compared. For each test, the subjects listened to 5 different tracks, each of which contained a speech signal time-scaled using one of the methods A, B, C, D1, or D2, and assigned the numbers 1 to 5 to each of the signals in order of their preference. The second part consisted of 8 A/B comparisons, in which the proposed method (D) was compared to the traditional uniform-scaling method (A). The subjects listened to two short tracks and selected one of the following ratings:  $A \ll B$  (A definitely worse than B),  $A < B$  (A slightly worse than B),  $A = B$  (no difference),  $A > B$  (A slightly better than B) and  $A \gg B$  (A definitely better than B). The results of the experiments show a clear preference for the proposed method, with 88% of listeners choosing a signal time-scaled by this method as their first choice in part one of the tests, see table 2.

Table 2. First Preference Allocations

Method:	First Preferences
D	88%
A or B or C	12%

Methods B and C show a small improvement in quality from that of A, but methods D1 and D2 lead the field by a much more significant amount as illustrated in the results shown in FIG. 7.



This pattern is noticeable for all time-scaling factors investigated, as can be seen from the results shown in Table 3.

Table 3. Preference Test Results

	SF 2	SF 2.5	SF 3
First	D2	D2	D2
Second	D1	D1	D1
Third	A	C	C
Fourth	C	B	B
Fifth	B	A	A

5

Also evident from Table 3 is the deterioration in quality of method A as the time-scaling factor increases. This can be observed more clearly from the results of the second part of the test, in which 78% of listeners chose method D over method A (table 4). The variation in this value with scaling factor forms the interesting result that, whereas method A decreases in quality as the scaling factor is increased, method D maintains a high quality output (as illustrated in FIG. 8).

10

Table 4. Comparison preferences

Method:	Preferences
D	78%
A	22%

15

The inventors of the present invention believe that, just as speech recognition systems perform best when trained on or adapted to the voice characteristics or speaking style of the speaker, the student's recognition of natural, colloquial speech improves more efficiently when trained on a similar natural speech corpus. Students will develop a higher level of recognition of faster speech signals, leading to the development of unambiguous perceptions of speech. The student's comprehension of the foreign language is enhanced using the methods of the present invention.

20

Although, the present invention has been described with reference to CALL packages, it may also be used with suitable modification in other applications

25

including voice mail messaging, talking books, web based audio browsing, broadcast radio, video replay with full audio at any speed, voice memo systems, transcription services, voice systems for the blind. Another application of the present invention would be as an assistive technology for use by speech therapists. In this context, a  
5 speech slow-down system (using the above described systems and methods) could be incorporated into a speech therapy system, so as to allow a speech therapist to slow speech so as to explain the generation of the sounds to a patient. This may be used by a speech therapist, for example, in the retraining of vocally impaired head-injured patients and/or stroke-victim patients, and/or in the speech training of special needs  
10 children. In this application, the system would be configured to allow a therapist to familiarise a patient with the characteristic sound of a well articulated phrase, and by slowing down or speeding up a recorded phrase, explain how the speech production mechanism is being controlled during its distinct phases. A small library of standard phrases could be prepared from which the associated characteristics could be  
15 determined manually. Other applications would include voice over matching / dubbing for film, audio forensics and web streaming. Accordingly, the present invention, although particularly advantageous in CALL packages, is not intended to be limited to these applications.

CLAIMS:

1. A voice playback system comprising an input means for accepting a segmented speech clip and a characteristic associated with each segment of the speech clip,  
5 time scale modification means for altering the duration of individual segments of the speech clip using a time-domain overlap-add technique to alter the duration of individual segments, wherein the level of alteration for the individual segment is determined with reference to the associated characteristic of the individual segment, and  
10 an output means for combining the altered individual segments into an altered duration speech clip, wherein the characteristic is selected from one or more of the following: a plosive, vowel, voiced consonant, unvoiced consonant or silence and where segments having a vowel characteristic are time scale modified at a rate of  $\alpha_1$ , segments excluding plosives having a voiced  
15 consonant characteristic are time scale modified at a rate of  $\alpha_2$ , and segments excluding plosives having a unvoiced consonant characteristic are time scale modified at a rate of  $\alpha_3$ , where  $\alpha_1 > \alpha_2 > \alpha_3$  when increasing the duration of the speech clip and/or  $\alpha_1 < \alpha_2 < \alpha_3$  when decreasing the duration of the speech clip.  
20
2. A voice playback system according to claim 1, wherein the time-domain overlap-add technique is a synchronised overlap-add algorithm.
3. A voice playback system according to claim 1, wherein the time-domain  
25 overlap-add technique is an adaptive overlap-add algorithm.
4. A voice playback system according to claim 1, wherein the level of alteration to be applied to individual segments is determined with reference to a user input associated with a desired playback speed.  
30
5. A voice playback system according to any preceding claim, wherein the duration of the segment is left unchanged when the associated characteristic is a plosive.

6. A voice playback system according to any preceding claim, wherein the level of alteration applied where the characteristic is a vowel is substantially equal to that applied where the characteristic is a silence.
- 5 7. A voice playback system according to any preceding claim, wherein  $\alpha_3 > 1$ .
8. A voice playback system according to any preceding claim wherein there is a substantially linear relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
- 10 9. A voice playback system according to any one of claims 1 to 11 wherein there is a substantially exponential relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
10. A computer assisted language learning system incorporating the voice playback system of any preceding claim.
- 15 11. A computer assisted language learning system according to claim 15, further comprising a computer readable storage medium having stored thereon at least one segmented speech clip comprising a plurality of segments, each segment having a characteristic associated with it, wherein the at least one characteristic identifies the suitability of the segment for time scale modification.
- 20 12. A language learning product comprising a datastore of altered duration speech produced by a system according to any preceding claim.
- 25 13. A method of altering the duration of a segmented speech clip comprising the steps of: altering the duration of individual segments of the speech clip using a time-domain overlap-add, wherein the level of alteration for the individual segment is determined with reference to an associated characteristic of the individual segment, and combining the altered individual segments into an altered duration speech clip, where the characteristic is selected from one or  
30 more of the following: a plosive, vowel, voiced consonant, unvoiced consonant or silence,  
wherein segments having a vowel characteristic are time scale modified at a rate

of  $\alpha_1$ , segments having a voiced consonant characteristic are time scale modified at a rate of  $\alpha_2$ , and segments having a unvoiced consonant characteristic are time scale modified at a rate of  $\alpha_3$ , where  $\alpha_1 > \alpha_2 > \alpha_3$  when increasing the duration of the speech clip and/or  $\alpha_1 < \alpha_2 < \alpha_3$  when decreasing the duration of the speech clip.

- 5
14. A method of altering the duration of a segmented speech clip according to claim 13, wherein a synchronised overlap-add algorithm technique is used for time scale modification.
- 10
15. A method of altering the duration of a segmented speech clip according to claim 14, wherein an adaptive overlap-add algorithm technique is used for time scale modification.
- 15
16. A method of altering the duration of a segmented speech clip according to any one of claims 13 to 15, wherein the level of alteration to be applied to individual segments is determined with reference to a user input associated with a desired playback speed.
- 20
17. A method of altering the duration of a segmented speech clip according to any of claims 13 to 16, wherein when the associated characteristic is a plosive the duration of the segment is left unchanged.
- 25
18. A method of altering the duration of a segmented speech clip according to any of claims 13 to 17, wherein the level of alteration where the characteristic is vowel is greater than the level of alteration where the characteristic is that of a voiced consonant.
- 30
19. A method of altering the duration of a segmented speech clip according to any of claims 13 to 18, wherein the level of alteration where the characteristic is a voiced consonant is greater than the level of alteration where the characteristic is an unvoiced consonant.

20. A method of altering the duration of a segmented speech clip according to any of claims 13 to 19, wherein the level of alteration where the characteristic is a voiced consonant is greater than the level of alteration where the characteristic is an unvoiced consonant.
- 5
21. A method of altering the duration of a segmented speech clip according to any of claims 13 to 20, wherein the level of alteration applied where the characteristic is a vowel is substantially equal to that applied where the characteristic is a silence.
- 10
22. A method of altering the duration of a segmented speech clip according to claim 13, wherein  $\alpha_3 > 1$ .
23. A method of altering the duration of a segmented speech clip according to claim 22, wherein there is a substantially linear relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
- 15
24. A method of altering the duration of a segmented speech clip according to claim 22, wherein there is a substantially exponential relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
- 20
25. The use of the method of any one of claims 13 to 24 in a computer assisted language learning environment.
- 25
26. A computer program product having code embodied therein which when run on a computer effects the running of a method according to any one of claims 13 to 24.
27. A computer program product according to claim 26, further comprising at least one segmented speech clip comprising a plurality of segments, each segment having a characteristic associated with it, wherein the at least one characteristic identifies the suitability of the segment for time scale modification.
- 30

28. A method of altering the duration of a speech clip comprising the steps of:  
segmenting the speech clip, associating a speech characteristic to each segment  
to produce a segmented speech clip and applying the method of any one of  
claims 13 to 27 to produce an altered duration speech clip.
- 5
29. A computer readable storage medium having stored thereon at least one  
segmented speech clip comprising a plurality of segments, each segment having  
associated characteristic stored with it.
- 10 30. A computer readable storage medium according to claim 29, wherein the  
characteristic is selected from one or more of the following: a plosive, vowel,  
voiced consonant, unvoiced consonant or silence.
- 15 31. A computer assisted language learning system comprising: a computer readable  
storage medium having stored thereon at least one segmented speech clip  
comprising a plurality of segments, each segment having a characteristic  
associated with it, wherein the at least one characteristic identifies the suitability  
of the segment for time scale modification,  
the computer assisted language learning system further comprising a voice  
20 playback means for playing the segmented speech clip to a user, wherein the  
voice playback means comprises time scale modification means for altering the  
duration of individual segments of the speech clip, wherein the level of  
alteration for the individual segment is determined with reference to the  
associated characteristic of the individual segment, and  
25 an output means for combining the altered individual segments into an altered  
duration speech clip for listening by the user.
- 30 32. A system according to claim 31, wherein the time scale modification means uses  
a time-domain overlap-add technique to alter the duration of individual  
segments.

33. A system according to claim 32, wherein the time-domain overlap-add technique is a synchronised overlap-add algorithm.
34. A system according to claim 32, wherein the time-domain overlap-add  
5 technique is an adaptive overlap-add algorithm.
35. A system according to claim 31, wherein the level of alteration to be applied to individual segments is determined with reference to a user input associated with a desired playback speed.  
10
36. A system according to any one of claims 31 to 35, wherein the characteristic is selected from one or more of the following: a plosive, vowel, voiced consonant, unvoiced consonant or silence.
- 15 37. A system according to any one of claims 31 to 36, wherein the duration of the segment is left unchanged when the associated characteristic is a plosive.
38. A system according to any one of claims 31 to 37, wherein the level of alteration where the characteristic is a vowel is greater than the level of alteration where  
20 the characteristic is that of a voiced consonant.
39. A system according to any one of claims 31 to 38, wherein the level of alteration where the characteristic is a voiced consonant is greater than the level of alteration where the characteristic is an unvoiced consonant.  
25
40. A system according to any one of claims 31 to 39, wherein the level of alteration applied where the characteristic is a vowel is substantially equal to that applied where the characteristic is a silence.
- 30 41. A system according to any one of claims 31 to 40, wherein segments having a vowel characteristic are time scale modified at a rate of  $\alpha_1$ , segments excluding plosives having a voiced consonant characteristic are time scale modified at a rate of  $\alpha_2$ , and segments excluding plosives having a unvoiced consonant



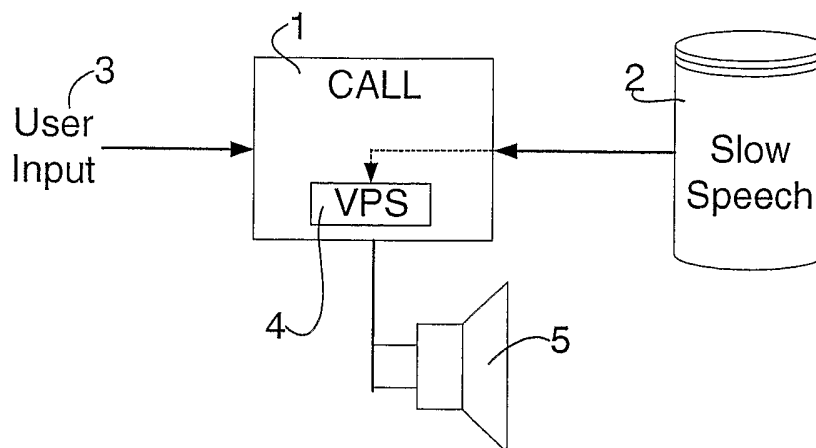
characteristic are time scale modified at a rate of  $\alpha_3$ , and where  $\alpha_1 > \alpha_2 > \alpha_3$ .

42. A system according to claim 41, wherein  $\alpha_3 > 1$ .
- 5 43. A system according to claim 41 or 42 wherein there is a substantially linear relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
44. A system according to claim 41 or 42 wherein there is a substantially exponential relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
- 10 45. A method of altering the duration of a segmented speech clip, comprising a plurality of segments, in a computer language learning system, the language learning system having an associated computer readable storage medium with the segmented speech clip stored thereon along with a characteristic associated with each segment, wherein the associated characteristic identifies the suitability of each individual segment for time scale modification, the method comprising the steps of: altering the duration of individual segments of the speech clip using time scale modification, wherein the level of alteration for the individual segment is determined with reference to the associated characteristic of the individual segment, and combining the altered individual segments into an altered duration speech clip.
- 15 46. A method according to claim 45, wherein a time-domain overlap-add technique is used for time scale modification.
- 20 47. A method according to claim 46, wherein a synchronised overlap-add algorithm technique is used for time scale modification.
- 25 48. A method according to claim 46, wherein an adaptive overlap-add algorithm technique is used for time scale modification.
- 30 49. A method according to any one of claims 45 to 48, wherein the level of alteration to be applied to individual segments is determined with reference to a

user input associated with a desired playback speed.

50. A method according to any of claims 45 to 49, wherein the characteristic is selected from one or more of the following: a plosive, vowel, voiced consonant, unvoiced consonant or silence.
51. A method according to any of claims 45 to 50, wherein when the associated characteristic is a plosive the duration of the segment is left unchanged.
52. A method according to any of claims 45 to 51, wherein the level of alteration where the characteristic is vowel is greater than the level of alteration where the characteristic is that of a voiced consonant.
53. A method according to any of claims 45 to 52, wherein the level of alteration where the characteristic is a voiced consonant is greater than the level of alteration where the characteristic is an unvoiced consonant.
54. A method according to any of claims 45 to 53, wherein the level of alteration where the characteristic is a voiced consonant is greater than the level of alteration where the characteristic is an unvoiced consonant.
55. A method according to any of claims 45 to 54, wherein the level of alteration applied where the characteristic is a vowel is substantially equal to that applied where the characteristic is a silence.
56. A method according to any of claims 45 to 55, wherein segments having a vowel characteristic are time scale modified at a rate of  $\alpha_1$ , segments having a voiced consonant characteristic are time scale modified at a rate of  $\alpha_2$ , and segments having a unvoiced consonant characteristic are time scale modified at a rate of  $\alpha_3$ , and where  $\alpha_1 > \alpha_2 > \alpha_3$ .
57. A method according to claim 56, wherein  $\alpha_3 > 1$ .

58. A method according to claim 56 or 57, wherein there is a substantially linear relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
59. A method according to claim 56 or 57, wherein there is a substantially exponential relationship between the values of  $\alpha_1$ ,  $\alpha_2$  and  $\alpha_3$ .
60. A voice playback system as described herein with reference to and/or as illustrated in the FIG. 2-8.
61. A computer assisted language learning system as described herein with reference to and/or as illustrated in the FIG. 2-8.
62. A method of altering the duration of a segmented speech clip as described herein with reference to and/or as illustrated in the FIG. 2-8.
63. A speech therapy system comprising a system of claims 1 to 10, wherein the speech therapy system further comprises a library of speech clips suitable for use by a speech therapist, each clip comprising a plurality of segments, each segment having a characteristic associated with it, wherein the at least one characteristic identifies the suitability of the segment for time scale modification.



Prior Art

FIG. 1

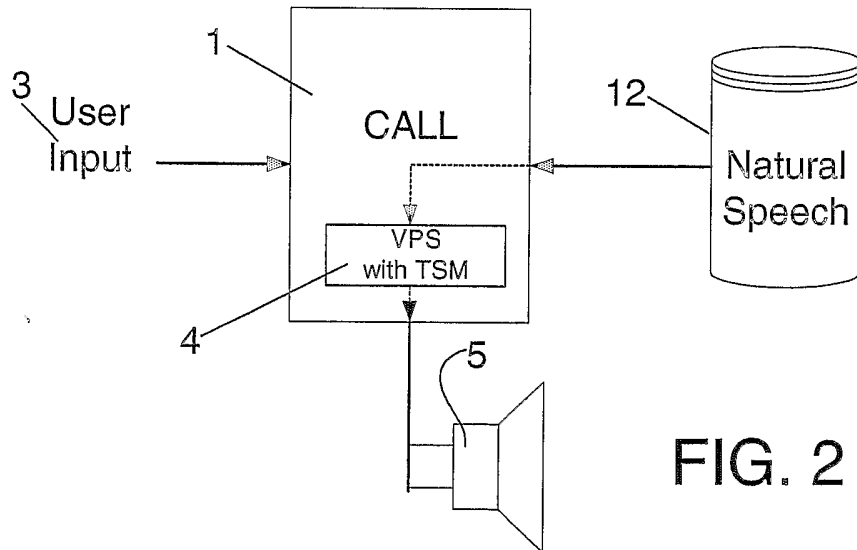


FIG. 2

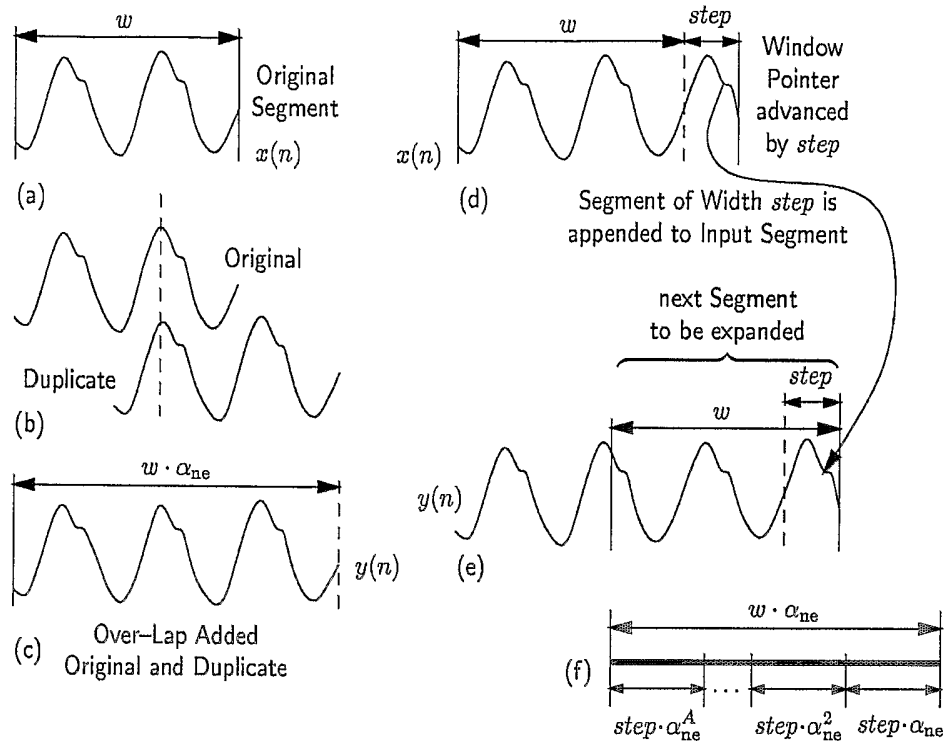


FIG. 3

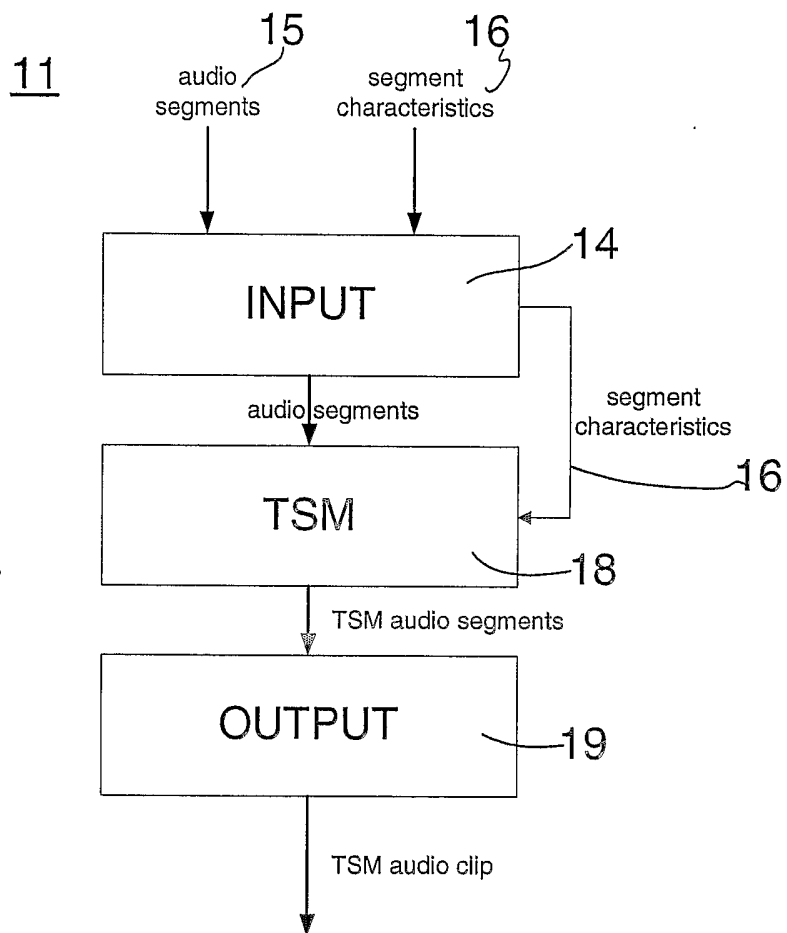


FIG. 4

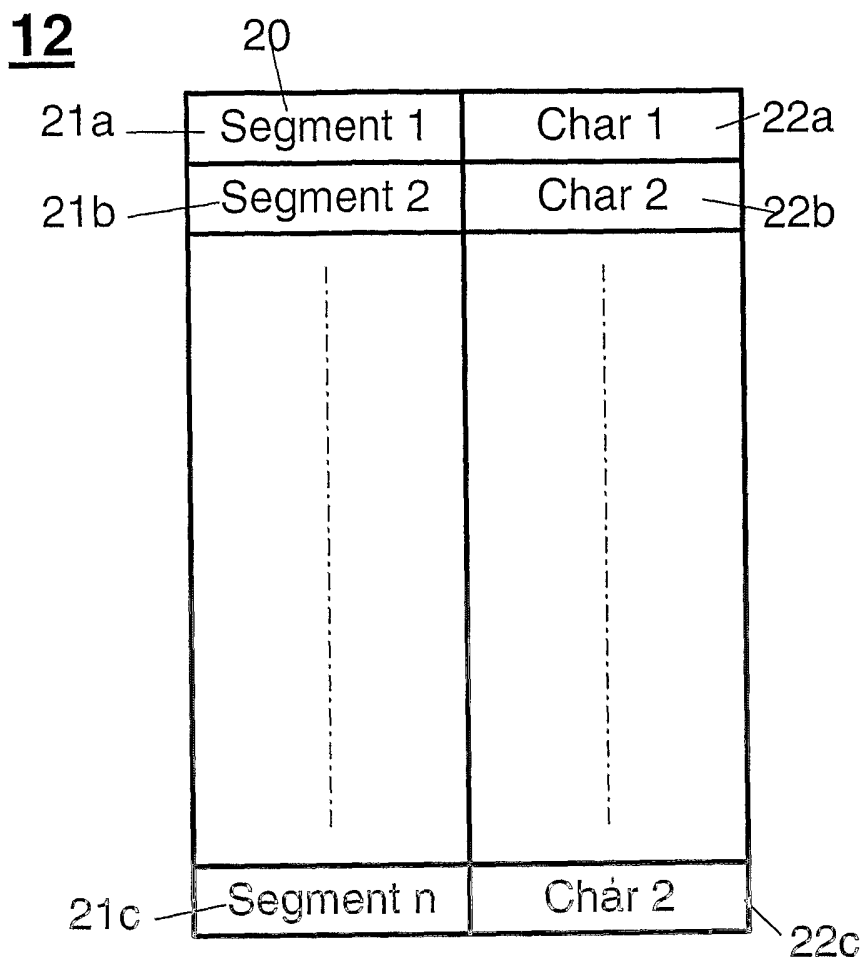


FIG. 5

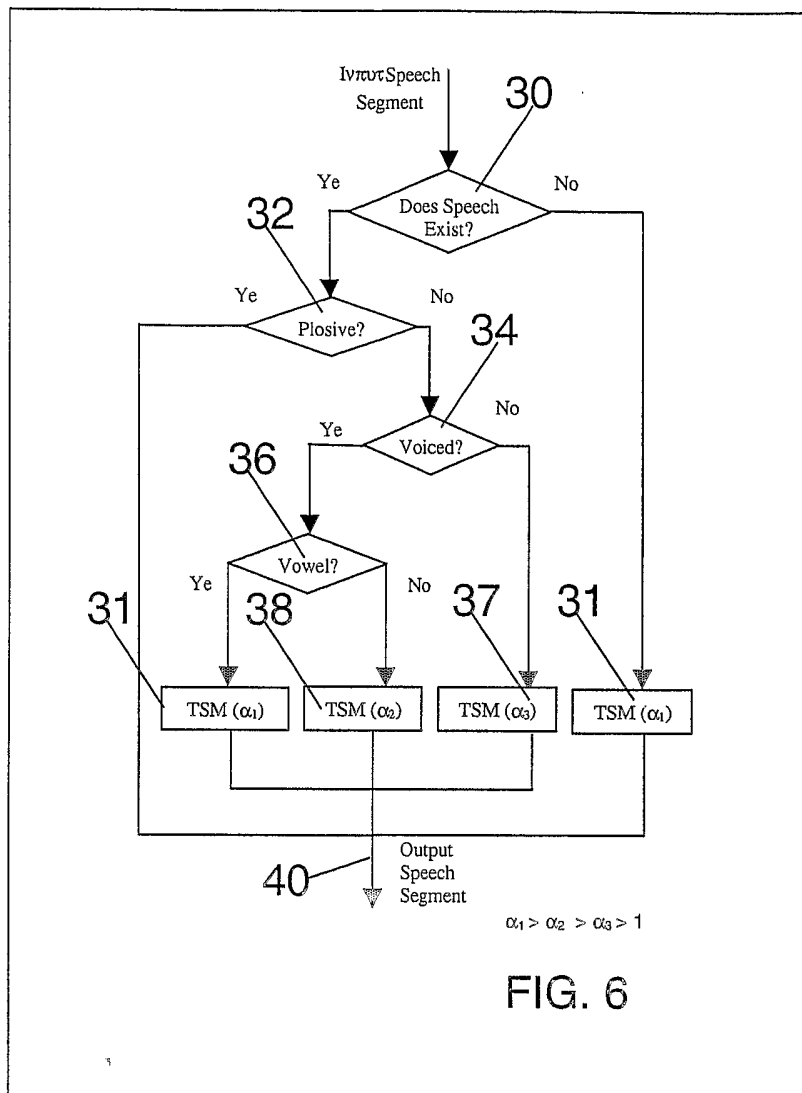


FIG. 6



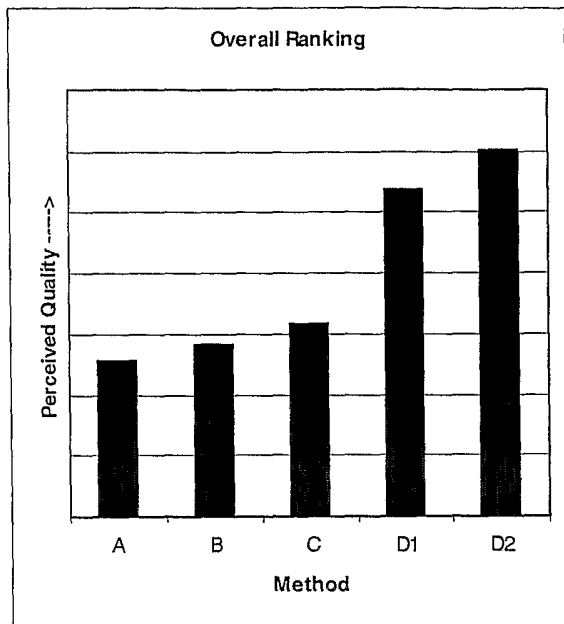


FIG. 7

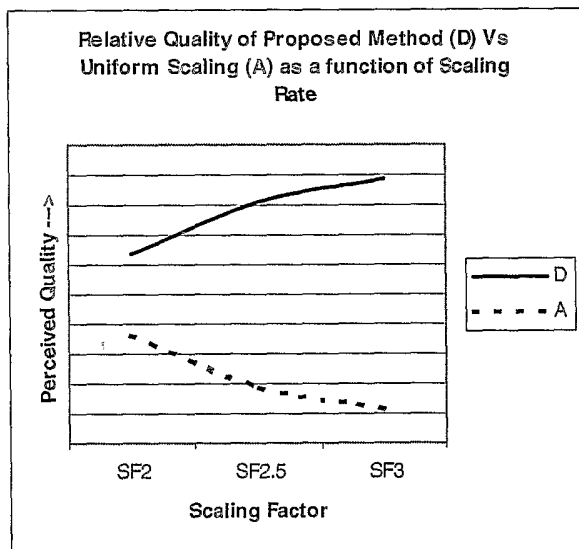


FIG. 8

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IE2004/000029

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 G09B19/04 G10L21/04 G10L13/02

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G09B G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WONG P H W ET AL: "On improving the intelligibility of synchronized over-lap-and-add (SOLA) at low TSM factor" TENCON '97. IEEE REGION 10 ANNUAL CONFERENCE. SPEECH AND IMAGE TECHNOLOGIES FOR COMPUTING AND TELECOMMUNICATIONS., PROCEEDINGS OF IEEE BRISBANE, QLD., AUSTRALIA 2-4 DEC. 1997, NEW YORK, NY, USA, IEEE, US, 2 December 1997 (1997-12-02), pages 487-490, XP010264299 ISBN: 0-7803-4365-4	1,2,4-9, 13,14, 16-24, 26,28-30
Y	abstract  * section 3. Modified SOLA for small TSM factor *  -----  -/--	3,10-12, 15,25, 27,31-63

Further documents are listed in the continuation of box C.

Patent family members are listed in annex.

° Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

1 June 2004

Date of mailing of the international search report

16/06/2004

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Quélavoine, R

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IE2004/000029

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	LAWLOR B ; FAGAN A D: "A novel efficient algorithm for music transposition" PROCEEDINGS 25TH EUROMICRO CONFERENCE, vol. 2, 8 September 1999 (1999-09-08), pages 48-54, XP002282755 MILAN, ITALY ISBN: 0-7695-0321-7 abstract; figure 3 -----	3,15,34, 48
Y	WO 98/44483 A (DIGISPEECH ISRAEL LTD ; SHAINSKI RINA (IL); SHPIRO ZEEV (IL)) 8 October 1998 (1998-10-08) abstract; figures 4b,7 -----	10-12, 25,27, 31-62
Y	US 6 413 098 B1 (JENKINS WILLIAM MICHAEL ET AL) 2 July 2002 (2002-07-02) abstract; figure 5 -----	63
X	SUNGJOO LEE ET AL: "Variable time-scale modification of speech using transient information" ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 1997. ICASSP-97., 1997 IEEE INTERNATIONAL CONFERENCE ON MUNICH, GERMANY 21-24 APRIL 1997, LOS ALAMITOS, CA, USA, IEEE COMPUT. SOC, US, 21 April 1997 (1997-04-21), pages 1319-1322, XP010226045 ISBN: 0-8186-7919-0 * section IV. Separating transient and steady portions for variable time-scale modification * abstract -----	1,2,4,5, 13,14, 16,17, 26,28, 29,60-62
P,X	DONNELLAN O ; JUNG E ; COYLE E: "Speech-adaptive time-scale modification for computer assisted language-learning" PROCEEDINGS 3RD IEEE INTERNATIONAL CONFERENCE ON ADVANCED TECHNOLOGIES, 9 July 2003 (2003-07-09), pages 165-169, XP010646630 ATHENS, GREECE the whole document -----	1-63

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No  
PCT/IE2004/000029

Patent document cited in search report	A	Publication date	Patent family member(s)	Publication date
WO 9844483	A	08-10-1998	AU 6516198 A WO 9844483 A1	22-10-1998 08-10-1998
<hr/>				
US 6413098	B1	02-07-2002	US 6302697 B1 US 6123548 A US 6413093 B1 US 6413094 B1 US 6413095 B1 US 6413096 B1 US 6413097 B1 AT 218002 T AU 712743 B2 AU 4408496 A CA 2206860 A1 DE 69526805 D1 DE 69526805 T2 EP 0797822 A1 JP 10511472 T WO 9618184 A1 US 6071123 A US 6413092 B1 US 5813862 A	16-10-2001 26-09-2000 02-07-2002 02-07-2002 02-07-2002 02-07-2002 02-07-2002 15-06-2002 18-11-1999 26-06-1996 13-06-1996 27-06-2002 07-11-2002 01-10-1997 04-11-1998 13-06-1996 06-06-2000 02-07-2002 29-09-1998
<hr/>				