

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

**EP 0 993 670 B1**

(12)

## EUROPEAN PATENT SPECIFICATION

(45) Date of publication and mention  
of the grant of the patent:

**20.03.2002 Bulletin 2002/12**

(21) Application number: **98932337.3**

(22) Date of filing: **01.07.1998**

(51) Int Cl.7: **G10L 21/02, G10L 13/02**

(86) International application number:  
**PCT/GB98/01936**

(87) International publication number:  
**WO 99/01863 (14.01.1999 Gazette 1999/02)**

### (54) **METHOD AND APPARATUS FOR SPEECH ENHANCEMENT IN A SPEECH COMMUNICATION SYSTEM**

VERFAHREN UND VORRICHTUNG ZUR SPRACHVERBESSERUNG IN EINEM  
SPRACHÜBERTRAGUNGSSYSTEM

PROCEDE ET APPAREIL D'AMELIORATION DE QUALITE DE SON VOCAL DANS UN SYSTEME  
DE COMMUNICATION PAR SON VOCAL

(84) Designated Contracting States:  
**AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE**  
Designated Extension States:  
**AL LT LV MK RO SI**

(30) Priority: **02.07.1997 GB 9714001**

(43) Date of publication of application:  
**19.04.2000 Bulletin 2000/16**

(73) Proprietor: **Simoco International Limited**  
**Cambridge CB4 1DP (GB)**

(72) Inventors:  
• **CHANCE, Robert James**  
**Birmingham B17 9SS (GB)**  
• **McLOUGHLIN, Ian Vince**  
**Birmingham B30 2RE (GB)**

(74) Representative: **Tothill, John Paul et al**  
**Frank B. Dehn & Co.**  
**179 Queen Victoria Street**  
**London EC4V 4EL (GB)**

(56) References cited:  
**EP-A- 0 732 686**

- **FANG-MING WANG ET AL: "FREQUENCY DOMAIN ADAPTIVE POSTFILTERING FOR ENHANCEMENT OF NOISY SPEECH" SPEECH COMMUNICATION, vol. 12, no. 1, 1 March 1993, pages 41-56, XP000382195**
- **BLAMEY P J ET AL: "FORMANT-BASED PROCESSING FOR HEARING AIDS" SPEECH COMMUNICATION, vol. 13, no. 3/04, 1 December 1993, pages 453-461, XP000421457**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

**EP 0 993 670 B1**

## Description

**[0001]** The present invention relates to a method and apparatus for speech enhancement in a speech communication system, and in particular to such a method and apparatus for enhancing speech to make it more intelligible to a listener in a noisy environment.

**[0002]** Speech communication systems such as mobile phones and radios are often used in noisy environments, such as inside vehicles. Furthermore, this environmental noise can vary during a conversation. This varying environmental noise can make it very difficult for a listener to understand the speech being output by their phone or radio.

**[0003]** EP-A-0732686, and the papers "Frequency Domain Adaptive Postfiltering for Enhancement of Noisy Speech", Wang et al, vol.12, no.1, March 1993, pages 41-56, Speech Communication, and "Formant-Based Processing for Hearing Aids", Blarney et al, vol. 13, no. 3/04 December 1993, pages 453-461, Speech Communication, all disclose methods of processing speech signals to reduce noise in the signals or to alter the speech signal in response to noise in the signal.

**[0004]** However, EP-A-0732686 merely discloses the use of an algorithm to map a speech signal into a given frequency range for transmission, the paper "Frequency Domain Adaptive Postfiltering for Enhancement of Noisy Speech" relates only to the suppression of noise in a speech signal and not to the alteration of the characteristics of speech, and the paper "Formant-Based Processing for Hearing Aids" relates to the modifications of speech signals but not in response to background noise.

**[0005]** According to one aspect of the present invention, there is provided a method for increasing the intelligibility of speech output by a speech communication system to a listener using the system, characterised by:

analysing the current background acoustic noise environment of the speech communication system; determining using the results of the background noise analysis whether the speech to be output to the listener would be intelligible to the listener in the current background noise environment by classifying the contents of the speech into at least two categories, and comparing the amplitude of the speech in one category at one frequency with the noise amplitude at that frequency; and altering the characteristics of the speech to be output by the speech communication system on the basis of said determination such that the altered speech output by the speech communication system has enhanced intelligibility to the listener in the current background noise.

**[0006]** According to a second aspect of the present invention, there is provided a speech communication system characterised by:

means for analysing the current background acoustic noise environment of the speech communication system;

means for determining using the results of the background noise analysis whether speech to be output by the speech communication system would be intelligible to a listener in the current background noise environment; and

means for altering the characteristics of the speech to be output by the speech communication system to enhance the intelligibility of the speech to a listener in the current background noise in accordance with the output of said determining means

wherein the means for determining whether the speech to be output would be intelligible comprises means for classifying the contents of the speech into different categories, and means for comparing the amplitude of one of the speech categories at one frequency with the noise amplitude at that frequency.

**[0007]** The present invention thus monitors the background noise in which a speech communication system is being used (i.e. the external environmental acoustic noise in the vicinity of the listener) and can adjust the characteristics of the speech to be output by the speech communication system to the listener to make it more intelligible in that current background acoustic noise. It therefore provides enhanced intelligibility of speech output as sound by, for example, the loudspeaker or ear-piece of a mobile phone or radio when used in noisy environments.

**[0008]** Furthermore, because the present invention analyses current background noise, it can take account of changes in the background noise and enhance the speech accordingly. In the present invention the background acoustic noise is therefore preferably continuously analysed and the speech continuously altered on the basis of that analysis. This provides for dynamic enhancement of the speech and is particularly advantageous in environments where background noise can change continuously and significantly, such as in a vehicle.

**[0009]** The background acoustic environmental noise can be analysed by various techniques, as is known in the art. It can be picked up or sampled using, for example, the usual microphone for picking up the user's speech of the speech communication system (e.g. mobile phone or radio), or a separate microphone.

**[0010]** An example background noise analysis system would be a process whereby the user's speech (for example in the microphone signal) is detected (using one of many common techniques, such as adding all input noise values in a given time interval and comparing these against a threshold) and the acoustic background noise is analysed during the gaps between the speech periods.

**[0011]** The sampled noise would then be analysed (perhaps using linear prediction) to determine both its

spectral content and its amplitude. LPC (linear prediction coefficient) values resulting from a linear predictive analysis contain sufficient spectral information, and a gain parameter could be used to relate the relative amplitudes of the LPC parameters to absolute amplitudes.

**[0012]** The intelligibility of speech to be output by the speech communication system in the current background noise is determined by classifying the contents of the speech into at least two categories, and comparing the amplitude of the speech in one category at one frequency with the noise amplitude at that frequency.

**[0013]** Preferably, descriptions of the speech and the background noise in the form of spectral analyses and amplitude scaling factor (gain) are compared to determine if the speech would be audible to a listener in that noise.

**[0014]** In one such comparison process, the speech contents could initially be classified into non-speech, voiced speech or unvoiced speech. If non-speech is present (perhaps a pause between words), then the audibility of this is unimportant and so it can be ignored.

**[0015]** If voiced speech is present, then its intelligibility needs to be determined. This is preferably done by comparing the amplitude of one or more, or most preferably each, spectral peak and/or of one or more, or most preferably each, formant (as is known in the art, voiced speech contains a series of resonant peaks at varying frequencies called formants which convey a great deal of information and to which spectral peaks in the spectral plot of the speech often correspond) in the voiced speech with the noise amplitude at the frequency of the peak or formant, respectively. If more than one peak or formant is to be considered, then the amplitude of each peak or formant should be compared with the noise amplitude at the frequency of the respective peak or formant.

**[0016]** Most preferably, the speech is determined to be unintelligible if the noise amplitude at any formant frequency or spectral peak or at a particular number of formant or spectral peak frequencies exceeds the corresponding formant or spectral peak amplitude(s).

**[0017]** Such comparison of the relative amplitudes of spectral peaks and formants in the speech with the background noise will give a good indication of the intelligibility of the speech, because it effectively determines the intelligibility of the speech in terms of a human listener model of intelligibility, i.e. it assesses the intelligibility of the speech in a manner that models closely a human listener's actual perception of the speech. As a well-known psycho-acoustic theory states, a sound of a given frequency will be masked by a second coincidental sound of similar frequency, and if the second sound is loud enough, then the former sound will be inaudible. Thus the Applicants have recognised that in the case of speech, loud noises with frequencies similar to those of formants or spectral peaks in the speech will mask the speech. Thus comparison of the amplitude of one or more or each formant or one or more or each spectral

peak in the speech with the noise amplitude at the corresponding frequency or frequencies will give a good indication of the audibility of that (or those) formant(s) or spectral peak(s) and thus of the intelligibility of the speech to a human listener.

**[0018]** Other speech classifications and categories could be used if desired. For example, the speech could be classified into vowel and consonant sounds (or other speech sounds). Preferably, a classification is used which is helpful or appropriate to determining intelligibility. Thus preferably, as in the above example, the classification includes a category which includes formants of the speech (preferably only formants) and that category is compared with the noise. Preferably the classification is into formant containing and non-formant containing categories.

**[0019]** Once the intelligibility of the speech has been determined, the speech can be altered to make it more intelligible in accordance with that determination. Preferably, if it is determined that the speech would be unintelligible, then the speech characteristics are altered, but not otherwise.

**[0020]** Alteration of the speech characteristics can be done in various ways, as is known in the art. It is preferably done by increasing the volume (amplitude) and/or altering the frequency of speech components and in particular the formants and/or spectral peaks in the speech.

**[0021]** In a particularly preferred such arrangement, the speech characteristics will be altered by adjusting the positions of the formants and/or spectral peaks in the speech spectral plot. Such alterations will have a more perceptible effect on the speech to a human listener and thus are particularly effective for increasing the intelligibility of the speech. For example, one or more peaks or formants could be shifted upwards or downwards in frequency, or the amplitude of one or more peaks or formants could be increased (corresponding to a decrease in bandwidth), or the bandwidth of one or more of the peaks or formants could be increased (corresponding to a decrease in amplitude).

**[0022]** Thus, for example, the volume of the formants can be increased such that they are audible over the background noise. However, this can be an undesirable way of altering the speech characteristics as speech volume levels sufficient to cause hearing loss (if sustained) may be required to make the speech intelligible in certain situations, notably those within noisy motor vehicles.

**[0023]** Preferably therefore the frequency of speech components such as formants or peaks in the speech spectrum is adjusted. This is preferably done to move them to a frequency where the noise level is lower, such that the components, e.g. peaks or formants, are audible (i.e. have an amplitude greater than the noise) at that frequency.

**[0024]** The alteration of speech characteristics is preferably carried out in accordance with the results of the

analysis of the background noise, and may be dependent upon the present or past values of the noise. Using present values of noise, a direct comparison may be made and an alteration made to the speech characteristics; using past values, it is possible to make predictive changes. For example, if the noise analysis indicates the noise amplitude reduces at a particular frequency to a level at which a presently inaudible formant would be audible, the speech characteristics could be altered to change the frequency of that formant to that particular frequency.

**[0025]** The actual alteration of speech characteristics can be carried out in a number of ways, as is known in the art. For example, the speech signal could be passed through an adaptive filter, such as a perceptual error weighting filter (as described in CHEN, J. H., COK, E. V., LIN, Y., JAYANT, N., and MIECHER, M.J., "A low delay CELP coder for the CCITT 16 kb/s speech coding standard". *IEEE J. Sci. Ateas Commun.* 1992, 10. (5). pp 830-849) to narrow or widen the formant bandwidth. Alternatively the amplitude peaks could be clipped so that the energy in the unvoiced parts of the speech becomes a more significant part of the total speech energy. This can increase intelligibility but at the expense of sound quality.

**[0026]** In a particularly preferred embodiment, the speech characteristics are altered by altering line spectral pair (LSP) data representing the speech.

**[0027]** As is known in the art, line spectral pairs are representations of the linear-prediction parameters derived for periods of sound. Where the sound is speech, the resonant frequencies in the speech or formants, can be noted in the linear-prediction spectrum. LSP values usually uniquely relate to positions of such resonances or formants in the linear-prediction spectrum. Thus LSP data can be used to represent speech, and the Applicants have recognised that by altering the LSP data, characteristics such as the frequency and amplitude of formants in the speech can be adjusted. This allows the speech characteristics to be adjusted relatively easily and in a way that can readily change the speech as perceived by a listener and at a much lower computational overhead than when using, for example, adaptive filtering. Also, such adjustment does not eliminate parts of the speech spectrum, but rather modifies them.

**[0028]** Furthermore, many speech communication systems such as speech coding/decoding systems used in mobile telephones or modern digital radio systems, utilise a linear-prediction model of speech, and convert this to an LSP representation for transmission. The LSP representation is generally used within such speech systems for reasons of information security and transmission efficiency.

**[0029]** Thus this embodiment of the present invention is particularly advantageous in such systems which use LSPs for speech transmission, since the LSP information that is transmitted may be altered in the speech communication system when it is received to enhance

the intelligibility of the speech. This altered LSP data would then be converted back to linear-prediction parameters and hence reconstructed into speech and output as sound, but with altered characteristics.

**[0030]** Preferably the frequency or the power and bandwidth of specific frequency-domain features, such as formants, found in the speech are altered in this way.

**[0031]** The LSP alterations can be designed to affect the reconstructed speech in specific ways so as to enhance the intelligibility of the speech over the background noise. For example, the particular line spectral pair (LSP) associated with a formant can be identified and its separation (or spacing) then widened or narrowed to increase or decrease the formant bandwidth. Alternatively or additionally, line spectral pairs can be moved higher or lower in frequency to increase or decrease the frequency of particular formants.

**[0032]** The LSP information is preferably altered by adding or subtracting values to one or more LSPs (or LSP lines), or by moving one or more LSPs (or LSP lines) in the speech spectrum. The values may be determined in accordance with the analysis of the background noise, and may be dependent upon the present or past values of each LSP. Using present values of LSP data, a direct comparison can be made with the ambient noise and an adjustment made to the LSP data; using past values, it is possible to make predictive changes.

**[0033]** In a particularly preferred such arrangement, the invention includes making a numerical increment or decrement in the value of any or all of the set of LSPs (or LSP lines) defining the speech. Thus individual or groups of LSPs can be moved to: shift one or more spectral peaks or formants in frequency (either upwards or downwards); or change the amplitude (either to increase the amplitude (decrease the bandwidth) or decrease the amplitude (increase the bandwidth)) of one or more spectral peaks or formants.

**[0034]** For example, the separation between the values of two or more of a set of LSP lines (and most preferably between a pair of LSP lines) can be narrowed or widened to narrow or widen frequency features (such as spectral peaks or formants) found in the speech frequency spectrum. Alternatively or additionally, the values of two or more of a set of LSP lines (and most preferably of a pair of LSP lines) can be incremented or decremented, most preferably by identical amounts (either in absolute terms or as a percentage of their original values), to adjust the centre frequency of features (such as spectral peaks or formants) found in the frequency spectrum of the speech.

**[0035]** In a particularly preferred embodiment, line spectral pairs are translated in frequency so as to change the centre frequency of particular peaks or formants in the speech data. As discussed above, this is a particularly advantageous way of changing speech characteristics as heard by a listener, for example to increase intelligibility over background noise.

**[0036]** It is also possible to predict the behaviour of

the background noise from an analysis of previous changes in its spectral content, to enable a faster or more appropriate adjustment to the LSPs. This is particularly applicable to repetitive noise such as a siren in a police car, fire appliance or ambulance. Knowledge of which way the frequency of the interfering noise is changing may affect the decision about which way to shift the formant frequencies.

**[0037]** Any or all of the above adjustments can be used individually or in combination to alter the speech characteristics of the speech to be output by the speech communication system in accordance with the analysis of the background noise of the listener to make the speech output by the speech communication system more intelligible to the listener.

**[0038]** The present invention has been described in relation to speech communication systems, such as mobile phones and radios. It is particularly suited to use in speech decoders, such as would be found for example in mobile phones or mobile radios. However, it would also be applicable (and in particular the aspects relating to LSP alteration would be applicable) to use in speech coders where it was desired to alter the characteristics of the user's input speech to be transmitted by the speech coder (for example to increase intelligibility over the speaker's background noise). It would also be applicable in radio receivers, televisions, or other devices which broadcast speech to listeners.

**[0039]** A preferred embodiment of the present invention will now be described by way of example only, and with reference to the accompanying drawings, in which:

Figure 1 shows a generic CELP codec structure;  
Figure 2 shows a block diagram of a typical speech communication system in accordance with the present invention;  
Figure 3 shows the frequency spectrum of a period of sound, with numbered LSP values for that sound overlaid as vertical lines; and  
Figure 4 shows the frequency spectrum of a period of sound derived from the LSP values of Figure 3 with specific alterations. The altered LSP values for that sound are overlaid as vertical lines.

**[0040]** The present invention is particularly applicable to use in a speech codec system such as would be used in a mobile phone or radio system. An example of such a codec structure is shown in Figure 1, in the form of a generic CELP coder.

**[0041]** The general CELP (codebook-excited linear prediction) structure was introduced in 1985 (see, for example, Shroeder MR, Atal BS, "Code-excited linear prediction (CELP): high-quality speech at very low bit rates", ICASSP, pp. 937-940, 1985), and many modifications have been made since.

**[0042]** A generic CELP codec structure 22 is shown in Figure 1. Figure 1 shows input speech 21 being analysed by linear prediction analyser unit or device 2 re-

sulting in linear prediction (LPC) parameters 3. The remainder of the input signal which linear prediction cannot describe is passed to a pitch filter, VQ encoding block 4 which produces parameters representative of, for example, the gain and pitch of the speech. These processes are unimportant to the invention and vary widely between different CELP implementations in their detail, however they result in various other parameters which, together with the LPC parameters, describe the input speech.

**[0043]** The LPC parameters 3 and any other parameters (such as gain and pitch) 5 describing the input speech are quantized by a quantizer 6 and transmitted (as transmission parameters 7) to the CELP decoder 14 which dequantizes them using a dequantizer 8. These dequantized values are then used to recreate speech 15 to be output as sound to a listener. (The dequantizer 8 reproduces the LPC parameters 3 and other parameters 5 by means of an LPC synthesiser 30 and pitch filter, VQ decoding block 31, respectively, which reproduce the speech for it to be output as sound 15.)

**[0044]** LPC parameters may alternatively be converted to a different form prior to quantization in the coder (and also converted back to LPC coefficients after dequantization). Such forms may include log area ratios, PARCOR (reflection coefficients) and line spectral pairs.

**[0045]** Differences in the representation of LPC parameter used and the types of (or usage of) pitch filter and vector quantizer (VQ) have led to many CELP variants. A small selection of examples are: MELP (mixed excitation linear prediction); VSELP (variable slope excitation linear prediction); SB-CELP (sub-band CELP); LD-CELP (low delay CELP); RELP (residual excitation linear prediction); RPE-LP (residual pulse excitation linear prediction); and others.

**[0046]** As noted above, in many such codecs the LPC parameters are transmitted as LSPs.

**[0047]** The terminology 'LSPs' refers to the parameters generated by a conversion of linear prediction coefficients using the line spectrum pair approach as described in the paper by Sugamura and Itakura (Sugamura N, Itakura F, "Speech analysis and synthesis methods developed at ECL in NTT - from LPC to LSP - ", Speech Communication, vol. 5, pp. 199-213, 1986). The linear prediction coefficients themselves are generated by any of the well-established analysis methods operating on a set of data (speech) such as those described in Makhoul J, "Linear prediction: a tutorial review", Proc. IEEE, vol 63, no. 4, pp. 561-580, 1975.

**[0048]** LSPs are generated via a mathematical transformation from LPCs and thus have identical information content, but different form. Many other mathematical transformations from LPCs have been determined, but none of the resulting parameters can be altered in the same way as LSPs and as described in the present invention.

**[0049]** The line spectral pair parameters may be referred to as line spectral frequencies, however this term

is not applied exclusively to LSPs.

**[0050]** Mathematically speaking, LSP parameters may be defined as: the roots of the two polynomials formed by a particular re-arrangement of the coefficients of the inverse linear prediction polynomial. These two polynomials may be called  $P$  and  $Q$  and are formed using the set of linear prediction coefficients,  $A_p$  (where  $p$  is the index of the array, usually running from 0 to the filter order,  $p$ ), having the following recursive relationship:

$$P(z^{-1}) = A_p(z^{-1}) - z^{-(p+1)} A_p(z)$$

$$Q(z^{-1}) = A_p(z^{-1}) + z^{-(p+1)} A_p(z)$$

The roots obtained by solving the polynomials  $P$  and  $Q$  give the line spectral frequency parameters, referred to as line spectral pairs. Many methods exist to determine these roots, as explained in, for example, the paper by Sugamura and Itakura referred to above. The choice of method is irrelevant for the purposes of the present invention.

**[0051]** The set of LSPs are often scaled. With reference to a 'basic' LSP value, the cosine or sine of these are also referred to as LSPs. In addition, the basic LSP may reside in one of various domains, i.e. its maximum and minimum values may be between 0 and  $\pi$ , between 0 and 4000Hz (a typical sampling frequency), or within other arbitrary ranges such as 0 to 1.

**[0052]** As an aid to understanding of the present invention, a non-mathematical description of line spectral pairs (LSPs) will also be considered. As LSPs are derived from LPC and reflection coefficients, it is necessary to cover these first.

**[0053]** Linear prediction is the usage of a fixed-length formula to model an unknown system. The formula structure is fixed but the values to be inserted into the formula must be found. Linear predictive analysis is the process of finding the best set of values for that formula. These values are the linear prediction coefficients, and the best set of these values is the set that causes the equation output to resemble the output of the system to be modelled most closely, when the inputs to the two systems are identical.

**[0054]** If the equation of that formula is re-ordered mathematically then another standard equation can be arrived at. The coefficients for the new equation are called reflection coefficients and can be found easily from the LPC coefficients.

**[0055]** The reflection coefficient equation is very easy to relate to a real system. For speech processing, the LPC analysis is attempting to find the best parameters that model a short period of speech. In physical terms, the model is made up of a number of different width but equal length tubes connected in series. The reflection coefficients fit well into this physical model as the reflection

coefficients relate directly to the difference between each consecutive tube.

**[0056]** When air is blown down tubes, resonances occur (organ pipes). In a human vocal tract, air originates at the glottis (which opens and closes rapidly) and proceeds through the vocal tract to be expelled at the mouth. The sound relates strongly to the shape of the vocal tract due to the resonances.

**[0057]** The LSP parameters each relate to the resonant frequency of one of the connected tubes. Half of the parameters are generated assuming that the source end of tubes is open, and half assuming that it is closed. In fact, the glottis opens and closes rapidly and so is neither open nor closed. Thus each true spectral resonance occurs between two nearby line spectral frequencies and these two values are considered to be a pair (thus line spectral pair).

**[0058]** An embodiment of the present invention in a speech communication system comprising a speech codec, and using LSP alteration to enhance the intelligibility of speech in a noisy environment is shown in Figure 2, and the signal processing is illustrated in Figures 3 and 4. The system as shown in Figure 2 has many features in common with the system of Figure 1 and thus the same reference numerals have been used for the like features of the systems.

**[0059]** The LSP alteration mechanism may act within a speech codec (a codec comprises both a coding 22 and a decoding 14 mechanism) in the positions shown in Figure 2 (i.e. in the speech decoder 14). The speech coder 22 transforms the input speech 21 into a set of condensed parameters 20 suitable for transmission by radio or other means to a receiving unit 14. (It should be noted that in this arrangement the LPC parameters produced by the linear prediction analyser 2 are converted to line spectral pair data by an LPC to LSP converter 32 before being quantized by the quantizer 6.) The receiving unit then decodes the transmitted data to reconstruct speech 15. By way of example, the coding unit 22 may reside in an office telephone and the decoding unit 14 within a mobile telephone handset.

**[0060]** In this embodiment alterations to the data received by the decoding unit, where that data comprises LSP information, are performed. This alteration unit is shown in Figure 2 as LSP processor 10.

**[0061]** The LSP processing depends upon the degree and type of acoustic noise background 16 that is present in the environment of the listener. The analysis unit 12 shown in Figure 2 determines the type and level of background noise by use of a microphone 13 which picks up, *inter alia*, the actual external background acoustic noise of the listener's environment.

**[0062]** An example of a noise analysis system would be a process whereby the user's speech is detected (using one of many common techniques, such as adding all input noise values in a given time interval and comparing these against a threshold) and the external acoustic background noise is considered during the

gaps between speech periods.

**[0063]** The sampled noise must then be analysed (perhaps using linear prediction) to determine both its spectral content and its amplitude. LPC (linear prediction coefficient) values resulting from a linear predictive analysis contain sufficient spectral information, and a gain parameter would relate the relative amplitudes of the LPC parameters to absolute amplitudes.

**[0064]** The decision device or unit 11 determines whether the speech data currently being received by the decoder and replayed as sound via the loudspeaker or ear piece of the mobile telephone unit would be intelligible to an average listener in the current background acoustic noise 16 of the mobile telephone unit (i.e. listener).

**[0065]** If the decision unit determines that speech is readily intelligible then no processing is necessary and the processing unit 10 would not alter the dequantized LSP parameters 17 which have been passed to it by the standard speech decoder, before passing them to the LSP to LPC converter 33.

**[0066]** On the other hand, if the decision unit determines that the speech is unintelligible, then processing is necessary and the processing unit 10 would alter the dequantized LSP parameters to alter the speech characteristics before passing them to the LSP to LPC converter for subsequent playback to the listener. The decision unit may also predict that the speech will shortly become unintelligible.

**[0067]** Inputs to the decision process are descriptions of speech and background noise, in the form of spectral analyses and amplitude scaling factor (gain). It is necessary to compare the speech and noise data to determine if the speech would be audible to a listener in that noise.

**[0068]** In this embodiment the comparison is done by initially classifying the contents of the speech signal into non-speech, voiced speech or unvoiced speech. If non-speech was present (perhaps a pause between words), then the audibility of this is unimportant and thus no enhancement is required, and the LSP-process module would be commanded to perform no processing.

**[0069]** If voiced speech is present (voiced speech contains a series of resonance peaks at various frequencies called formants), then the amplitude of each formant would be compared to the noise amplitude at that frequency to determine its audibility. If the noise amplitude at any formant frequency exceeds the formant amplitude then formant adjustment is required.

**[0070]** The LSP process unit 10 performs mathematical operations on individual LSPs to enhance the speech under the control of the decision unit.

**[0071]** The exact operations would depend upon the directions of the decision process. One speech enhancement function would entail the shifting of LSP lines to more favourable locations.

**[0072]** For example, an automatic examination of the noise amplitudes around the formant frequency might

reveal if, perhaps, shifting the formant frequency upwards or downwards by 10% may improve matters. If this is likely (perhaps because the noise amplitude reduces at a frequency 10% lower than the formant frequency), then the LSP processing block is directed to shift the appropriate LSPs by the corresponding amount.

**[0073]** If, for example, the formant that requires moving is located at 600Hz, then two LSP coefficients would exist, usually very close to and either side of 600Hz. If audibility is to be improved by a downwards shift of 10%, then the values of these two LSP parameters would each be multiplied by 0.9 to effect that shift. The LSP adjustment itself is confined to within the LSP process block.

**[0074]** As a further example, if the decision module determined that shifting lines 1 and 2 from a set of LSPs downwards in frequency by 10% would improve intelligibility, then the values of lines 1 and 2 would both be multiplied by a factor of 0.9.

**[0075]** If the decision module determined that upward shifting of line 3 by 100Hz improves intelligibility then an amount would be added to line 3. This amount would be equal to 100 if the LSP parameters were scaled to have values in Hz, or would more generally be

$$\frac{100 \times 2\pi}{f_s}$$

where  $f_s$  is the sampling rate of the system, and the values of the LSPs are confined to the angular frequency domain.

**[0076]** Other types of processing are possible, but may all be described as adding/subtracting values to one or more LSP lines (with adding LSP lines to themselves being equivalent to multiplication). The values may be determined by the decision module or may be dependent upon the present or past value of each LSP line.

**[0077]** An example of such LSP processing is illustrated in Figure 3, in which the frequency spectrum of a period of sound has been plotted, and the 10 LSP lines obtained from analysing this sound have been overlaid. LSP values may be readily converted to and from the LPC parameters from which the spectrum is plotted. For the specific example in question, Figure 3 thus shows the frequency spectrum of the sound obtained from the analysis of speech 21 in the CELP coder 22 of Figure 2.

**[0078]** In the case of a standard CELP decoder, operating without the benefit of this invention, the output speech 15 would be reconstructed using the data of Figure 3. When the invention is included, the LSP processing block 10 would be capable of altering the LSP values in order to change the output speech 15.

**[0079]** For the specific example of Figure 4, certain of the LSP values of the spectrum of Figure 3 have been altered and a new set of LPC coefficients have thus

been generated forming the spectrum as shown in Figure 4. Referring to the LSP values of the original spectrum in Figure 3, three operations have been performed:

1. The separation between lines 1 and 2 has been increased by moving both of the lines further apart (in other words 1 has been lowered in frequency and 2 has been raised)
2. Lines 5 and 6 have been increased in frequency
3. Line 10 has been increased in frequency.

The three actions have specific consequences to the sound that is transmitted:

1. Lines 1 and 2 lie on either side of a spectral peak. The movement in the two lines has induced this spectral peak to both reduce in amplitude and become wider (equivalent to an increase in bandwidth).
2. Lines 5 and 6 lie on either side of a second spectral peak. The movement of these two lines has induced that peak to increase in frequency.
3. Line 10 previously lay to the right of a very small spectral 'bump' which is now no longer evident as the line has been increased in frequency by a substantial amount.

**[0080]** In this specific example of a speech codec, the sound under analysis is speech. The spectral peaks evident in the spectral plots will then often, as discussed above, correspond to formants, important constituents of speech that convey a great deal of information. The LSP-based adjustments discussed above have thus changed the characteristics of the speech to be output to and as it will be perceived by the listener. For example, in the case of vowels, moderately widening the lines corresponding to spectral peaks (i.e. increasing the bandwidths of the formants) has been found to improve intelligibility.

**[0081]** The example shown in Figure 2 additionally analyses the noise present in the environment of the listener to determine if the speech to be replayed to that listener is intelligible. If not, then speech characteristics are altered in the present invention to improve the intelligibility of the speech by the operation of moving individual or groups of LSPs to provide the following set of operations:

1. Shift peak/formant upwards in frequency.
2. Shift peak/formant downwards in frequency.
3. Increase amplitude (decrease bandwidth) of peak/formant.

4. Increase bandwidth (decrease amplitude) of peak/formant.

**[0082]** A well-known psychoacoustic theory states that a sound of given frequency will be masked by a second coincidental sound of similar frequency. If the second sound is loud enough, then the former sound will be inaudible. Thus, in the case of speech, the Applicants have recognised that loud noises with frequencies similar to those of the formants will mask the speech. In order to hear the speech it is necessary to either increase the volume or alter the frequency of the speech components.

**[0083]** Volume alteration is relatively straightforward, but it should be noted that speech volume levels sufficient to cause hearing loss (if sustained) may be required to make speech intelligible in certain situations, notably those within noisy motor vehicles. It is therefore preferred to alter the frequency of speech components.

**[0084]** As can be seen, the present invention offers a method of reducing the masking of speech by acoustic background noise (and thus improving intelligibility) through an efficient process that may be combined with many of the current standard mobile telephone and radio systems, and standard speech codecs in such systems.

**[0085]** Speech enhancement results when an analysis of the listener's background noise environment is combined with corrective LSP alteration, which adjusts received transmitted speech data to be replayed to the listener in order to improve the chances of the listener hearing the processed sounds. The technique adjusts the values of LSPs found within the speech data codec based upon an analysis of the background acoustic noise environment of the listener. Preferably, the frequency or the power and bandwidth of specific frequency-domain features found in the received speech are altered in this way.

## Claims

1. A method for increasing the intelligibility of speech output by a speech communication system to a listener using the system, comprising:

analysing the current background acoustic noise environment of the listener;  
determining using the results of the background noise analysis whether the speech to be output to the listener would be intelligible to the listener in their current background noise environment by classifying the contents of the speech into at least two categories, and comparing the amplitude of the speech in one category at one frequency with the noise amplitude at that frequency; and  
altering the characteristics of the speech to be



output by the speech communication system on the basis of said determination such that the altered speech has enhanced intelligibility to the listener in their current background noise environment.

2. A method as claimed in claim 1, wherein the intelligibility of the speech to be output is determined by classifying the contents of the speech into a category which contains formants in the speech, and comparing the amplitude of the formant containing speech category at one frequency with the noise amplitude at that frequency. 10
3. A method as claimed in claim 1 or claim 2, wherein the intelligibility of the speech to be output is determined by classifying the contents of the speech into non-speech, voiced speech or unvoiced speech, and comparing the amplitude of the voiced speech at one frequency with the noise amplitude at that frequency. 15 20
4. A method as claimed in any one of claims 1 to 3, wherein the intelligibility of the speech to be output is determined by classifying the contents of the speech into non-speech, voiced speech or unvoiced speech, and comparing the amplitude of a spectral peak of the voiced speech having a centre frequency, with the noise amplitude at the centre frequency of the spectral peak. 25 30
5. A method as claimed in any one of claims 1 to 4, wherein the intelligibility of the speech to be output is determined by classifying the contents of the speech into non-speech, voiced speech or unvoiced speech, and comparing the amplitude of a formant of the voiced speech having a centre frequency, with the noise amplitude at the centre frequency of the formant. 35 40
6. A method as claimed in any of claims 1 to 5, wherein the speech is determined to be unintelligible if the background noise amplitude at substantially the same frequency as a spectral peak in the speech exceeds the amplitude of the spectral peak. 45
7. A method as claimed in any one of claims 1 to 6, wherein the speech is determined to be unintelligible if the background noise amplitude at substantially the same frequency as a formant in the speech exceeds the amplitude of the formant. 50
8. A method as claimed in any one of claims 1 to 7, wherein the speech characteristics are altered by altering line spectral pair (LSP) data representing the speech. 55
9. A method as claimed in claim 8, wherein the speech

characteristics are altered by moving a line spectral pair in the speech spectrum.

10. A method as claimed in any one of claims 1 to 9, wherein the speech characteristics are altered by altering the frequency of a component in the speech spectrum.
11. A method as claimed in claim 10, wherein the frequency of a formant in the speech spectrum is altered.
12. A method as claimed in claim 11, wherein the frequency of a formant in the speech is altered to move the formant to a frequency where the background noise amplitude is lower.
13. A method as claimed in any one of claims 10 to 12, wherein the speech spectrum includes a spectral peak having a centre frequency, and the centre frequency of the spectral peak in the speech spectrum is altered.
14. A speech communication system comprising:
  - means (12) for analysing the current background acoustic noise environment of the speech communication system;
  - means (11) for determining using the results of the background noise analysis whether speech to be output by the speech communication system to a listener listening to the speech communication system would be intelligible to the listener in the current background noise environment; and
  - means (10) for altering the characteristics of the speech to be output by the speech communication system to the listener to enhance the intelligibility of the speech to the listener in the current background noise in accordance with the output of said determining means, wherein the means (11) for determining whether the speech to be output would be intelligible comprises means for classifying the contents of the speech into different categories, and means for comparing the amplitude of one of the speech categories at one frequency with the noise amplitude at that frequency.
15. A system as claimed in claim 14, wherein the means for classifying the contents of the speech into different categories classifies the contents of the speech into a category which contains formants in the speech, and the comparing means compares the amplitude of the formant containing speech category at one frequency with the noise amplitude at that frequency.

16. A system as claimed in claim 14 or claim 15, wherein the means (11) for determining whether the speech to be output would be intelligible comprises means for comparing the noise amplitude at substantially the same frequency as a formant in the speech with the amplitude of the formant. 5
17. A system as claimed in any one of claims 14 to 16, wherein the speech is represented by data including line spectral pair (LSP) data, and the means (10) for altering the characteristics of the speech to be output by the speech communication system comprises means for altering the line spectral pair (LSP) data representing the speech. 10
18. A system as claimed in any one of claims 14 to 17, wherein the means (10) for altering the characteristics of the speech to be output by the speech communication system comprises means for altering the frequency of a component in the speech spectrum. 15
19. A system as claimed in claim 18, wherein the means (10) for altering the characteristics of the speech to be output by the speech communication system comprises means for altering the frequency of a formant in the speech to move the formant to a frequency where the noise amplitude is lower. 20

#### Patentansprüche

1. Verfahren zum Verbessern bzw. Steigern der Verständlichkeit einer Sprachausgabe durch ein Sprachkommunikationssystem für einen das System nutzenden Hörer mit folgenden Verfahrensschritten: 25

Analysieren der gegenwärtigen akustischen Hintergrund-Rauschumgebung des Hörers, 40

Bestimmen unter Verwendung der Ergebnisse der Hintergrund-Rauschanalyse, ob die zu dem Hörer auszugebende Sprache für den Hörer in seiner gegenwärtigen Hintergrund-Rauschumgebung verständlich sein würde, indem die Inhalte der Sprache in mindestens zwei Kategorien eingestuft werden, und die Amplitude der Sprache in einer Kategorie bei einer Frequenz mit der Rauschamplitude bei dieser Frequenz verglichen wird, und 50

Ändern der Eigenschaften bzw. Kenndaten der Sprache, die durch das Sprachkommunikationssystem auszugeben ist, auf Basis der Bestimmung, so daß die veränderte Sprache eine verbesserte Verständlichkeit für den Hörer in seiner gegenwärtigen Hintergrund-Rauschum-

gebung hat.

2. Verfahren nach Anspruch 1, bei dem die Verständlichkeit der auszugebenden Sprache bestimmt wird durch Einstufen des Inhalts der Sprache in eine Kategorie, die Formance in der Sprache enthält, und durch Vergleichen der Amplitude der Sprachkategorie, die Formante enthält, bei einer Frequenz mit der Rauschamplitude bei dieser Frequenz.
3. Verfahren nach Anspruch 1 oder 2, bei dem die Verständlichkeit der auszugebenden Sprache bestimmt wird durch Einstufen des Inhalts der Sprache in Nicht-Sprache, gesprochene Sprache oder ungesprochene Sprache, und durch Vergleichen der Amplitude der gesprochenen Sprache bei einer Frequenz mit der Rauschamplitude bei dieser Frequenz.
4. Verfahren nach einem der Ansprüche 1 bis 3, bei dem die Verständlichkeit der auszugebenden Sprache bestimmt wird durch Einstufen des Inhalts der Sprache in Nicht-Sprache, gesprochene Sprache oder ungesprochene Sprache, und durch Vergleichen der Amplitude einer spektralen Spitze der gesprochenen Sprache, die eine Mittenfrequenz bzw. Ruhefrequenz hat, mit der Rauschamplitude bei der Mittenfrequenz der spektralen Spitze.
5. Verfahren nach einem der Ansprüche 1 bis 4, bei dem die Verständlichkeit der auszugebenden Sprache bestimmt wird durch Einstufen des Inhalts der Sprache in Nicht-Sprache, gesprochene Sprache oder ungesprochene Sprache, und durch Vergleichen der Amplitude eines Formants der gesprochenen Sprache, die eine Mittenfrequenz hat, mit der Rauschamplitude bei der Mittenfrequenz des Formants. 30
6. Verfahren nach einem der Ansprüche 1 bis 5, bei dem die Sprache als unverständlich bestimmt wird, wenn die Hintergrund-Rauschamplitude bei im wesentlichen der gleichen Frequenz wie eine spektrale Spitze in der Sprache die Amplitude der spektralen Spitze übertrifft. 35
7. Verfahren nach einem der Ansprüche 1 bis 6, bei dem die Sprache als unverständlich bestimmt wird, wenn die Hintergrund-Rauschamplitude bei im wesentlichen der gleichen Frequenz wie ein Formant in der Sprache die Amplitude des Formants übertrifft.
8. Verfahren nach einem der Ansprüche 1 bis 7, bei dem die Spracheigenschaften bzw. Sprachkenndaten durch Ändern der Daten eines Spektrallinienpaares (line spectral pair: LSP), die die Sprache repräsentieren, geändert werden. 40 45 50 55

9. Verfahren nach Anspruch 8, bei dem die Spracheigenschaften durch Bewegen eines Spektrallinienspaars in das Sprachspektrum geändert werden.
10. Verfahren nach einem der Ansprüche 1 bis 9, bei dem die Spracheigenschaften durch Ändern der Frequenz einer Komponente in dem Sprachspektrum geändert werden. 5
11. Verfahren nach Anspruch 10, bei dem die Frequenz eines Formants in dem Sprachspektrum geändert wird. 10
12. Verfahren nach Anspruch 11, bei dem die Frequenz eines Formants in der Sprache geändert wird, um den Formant zu einer Frequenz zu bewegen, bei der die Hintergrund-Rauschamplitude niedriger ist. 15
13. Verfahren nach einem der Ansprüche 10 bis 12, bei dem das Sprachspektrum eine spektrale Spitze mit einer Mittenfrequenz umfaßt und die Mittenfrequenz der spektralen spitze in dem Sprachspektrum geändert wird. 20
14. Sprachkommunikationssystem mit: 25
- Mitteln (12) zum Analysieren der gegenwärtigen akustischen Hintergrund-Rauschumgebung des Sprachkommunikationssystems,
- Mitteln (11) zum Bestimmen unter Verwendung der Ergebnisse der Hintergrund-Rauschanalyse, ob die durch das Sprachkommunikationssystem an einen Hörer, der dem Sprachkommunikationssystem zuhört, auszugebende Sprache, verständlich für den Zuhörer in der gegenwärtigen Hintergrund-Rauschumgebung sein würde, und 30
- Mitteln (10) zum Ändern der Eigenschaften der durch das Sprachkommunikationssystem für den Hörer auszugebenden Sprache, um die Verständlichkeit der Sprache für den Hörer in dem gegenwärtigen Hintergrundrauschen zu verbessern, gemäß der Ausgabe der Bestimmungsmittel, 35
- bei dem das Mittel (11) zum Bestimmen, ob die auszugebende Sprache verständlich sein würde, Mittel zum Einstufen des Inhalts der Sprache in unterschiedliche Kategorien umfaßt, und Mittel zum Vergleichen der Amplitude einer Sprachkategorie bei einer Frequenz mit der Rauschamplitude bei dieser Frequenz. 40
15. System nach Anspruch 14, bei dem das Mittel zum Einstufen des Inhalts der Sprache in unterschiedliche Kategorien den Inhalt der Sprache in eine Kategorie einstuft, die Formanten in der Sprache enthält, und das Vergleichsmittel die Amplitude der Sprachkategorie, die Formanten enthält, bei einer Frequenz mit der Rauschamplitude bei dieser Frequenz vergleicht. 45
16. System nach Anspruch 14 oder 15, bei dem das Mittel (11) zum Bestimmen, ob die auszugebende Sprache verständlich sein würde, Mittel zum Vergleichen der Rauschamplitude bei der im wesentlichen gleichen Frequenz wie ein Formant in der Sprache mit der Amplitude des Formante umfaßt.
17. System nach einem der Ansprüche 14 bis 16, bei dem die Sprache durch Daten repräsentiert ist, die Daten eines Spektrallinienspaars (LSP) enthalten, und das Mittel (10) zum Ändern der Eigenschaften der durch das Sprachkommunikationssystem auszugebenden Sprache Mittel zum Ändern der Daten des Spektrallinienspaars (LSP) umfaßt, die die Sprache repräsentieren.
18. System nach einem der Ansprüche 14 bis 17, bei dem das Mittel (10) zum Ändern der Eigenschaften der durch das Sprachkommunikationssystem auszugebenden Sprache Mittel zum Ändern der Frequenz einer Komponente in dem Sprachspektrum umfaßt.
19. System nach Anspruch 18, bei dem das Mittel (10) zum Ändern der Eigenschaften der durch das Sprachkommunikationssystem auszugebenden Sprache Mittel zum Ändern der Frequenz eines Formants in der Sprache, um den Formant zu einer Frequenz zu bewegen, bei der die Rauschamplitude niedriger ist, umfaßt.

## Revendications

1. Procédé qui permet d'accroître l'intelligibilité d'un signal vocal délivré par un système de communication vocal à un auditeur qui utilise le système, comprenant les étapes suivantes :
- analyser l'environnement du bruit de fond acoustique actuel de l'auditeur ;
- déterminer, en utilisant les résultats de l'analyse du bruit de fond, si le signal vocal à délivrer à l'auditeur serait intelligible pour l'auditeur dans son environnement de bruit de fond actuel en classant le contenu du signal vocal en deux catégories au moins, et en comparant l'amplitude du signal vocal, dans une catégorie à une fréquence, à l'amplitude du bruit à cette fréquence ; et
- modifier les caractéristiques du signal vocal qui doit être délivré par le système de communica-

tion vocal sur la base de ladite détermination de sorte que le signal vocal modifié ait une intelligibilité améliorée pour l'auditeur dans son environnement de bruit de fond actuel.

2. Procédé selon la revendication 1, dans lequel l'intelligibilité du signal vocal qui doit être délivré est déterminée en classant le contenu du signal vocal dans une catégorie qui contient des formants dans le signal vocal, et en comparant l'amplitude du formant qui contient la catégorie du signal vocal à une fréquence, à l'amplitude du bruit à cette fréquence. 5
3. Procédé selon l'une quelconque des revendications 1 ou 2, dans lequel l'intelligibilité du signal vocal qui doit être délivré est déterminée en classant le contenu du signal vocal en signal non-vocal, en signal voisé ou en signal non-voisé, et en comparant l'amplitude du signal voisé à une fréquence, à l'amplitude du bruit à cette fréquence. 10
4. Procédé selon l'une quelconque des revendications 1 à 3, dans lequel l'intelligibilité du signal vocal qui doit être délivré est déterminée en classant le contenu du signal vocal en signal non-vocal, en signal voisé ou en signal non-voisé, et en comparant l'amplitude d'une pointe spectrale du signal voisé ayant une fréquence centrale, à l'amplitude du bruit à la fréquence centrale de la pointe spectrale. 15
5. Procédé selon l'une quelconque des revendications 1 à 4, dans lequel l'intelligibilité du signal vocal qui doit être délivré est déterminée en classant le contenu du signal vocal en signal non-vocal, en signal voisé ou en signal non-voisé, et en comparant l'amplitude d'un formant du signal voisé ayant une fréquence centrale, à l'amplitude du bruit à la fréquence centrale du formant. 20
6. Procédé selon l'une quelconque des revendications 1 à 5, dans lequel le signal vocal est déterminé comme étant inintelligible si l'amplitude du bruit de fond à sensiblement la même fréquence qu'une pointe spectrale du signal vocal, excède l'amplitude de la pointe spectrale. 25
7. Procédé selon l'une quelconque des revendications 1 à 6, dans lequel le signal vocal est déterminé comme étant inintelligible si l'amplitude du bruit de fond à sensiblement la même fréquence qu'un formant du signal vocal, excède l'amplitude du formant. 30
8. Procédé selon l'une quelconque des revendications 1 à 7, dans lequel les caractéristiques du signal vocal sont modifiées par la modification des données de paire de raies spectrales (ou LSP, acronyme de Line Spectral Pair) qui représentent le signal vocal. 35

9. Procédé selon la revendication 8, dans lequel les caractéristiques du signal vocal sont modifiées par le déplacement d'une paire de raies spectrales dans le spectre du signal vocal.

10. Procédé selon l'une quelconque des revendications 1 à 9, dans lequel les caractéristiques du signal vocal sont modifiées par la modification de la fréquence d'une composante du spectre du signal vocal.

11. Procédé selon la revendication 10, dans lequel la fréquence d'un formant du spectre du signal vocal est modifiée.

12. Procédé selon la revendication 11, dans lequel la fréquence d'un formant du spectre du signal vocal est modifiée afin de déplacer le formant à une fréquence où l'amplitude du bruit de fond est plus faible.

13. Procédé selon l'une quelconque des revendications 10 à 12, dans lequel le spectre du signal vocal comprend une pointe spectrale ayant une fréquence centrale, et la fréquence centrale de la pointe spectrale du spectre du signal vocal est modifiée.

14. Système de communication vocal comprenant :

des moyens (12) pour analyser l'environnement du bruit de fond acoustique actuel du système de communication vocal ;  
des moyens (11) pour déterminer, en utilisant les résultats de l'analyse du bruit de fond, si le signal vocal qui doit être délivré par le système de communication vocal à un auditeur qui écoute le système de communication vocal serait intelligible pour l'auditeur dans l'environnement de bruit de fond actuel ; et  
des moyens (10) pour modifier les caractéristiques du signal vocal qui doit être délivré par le système de communication par signal vocal à un auditeur afin d'améliorer l'intelligibilité du signal vocal pour l'auditeur dans l'environnement de bruit de fond actuel conformément à la sortie desdits moyens de détermination,

dans lequel les moyens (11) pour déterminer si le signal vocal qui doit être délivré serait intelligible, comprennent des moyens qui permettent de classer le contenu du signal vocal en différentes catégories, et des moyens qui permettent de comparer l'amplitude de l'une des catégories du signal vocal à une fréquence, à l'amplitude du bruit à cette fréquence.

15. Procédé selon la revendication 14, dans lequel les moyens qui permettent de classer le contenu du signal vocal en différentes catégories classent le con-

tenu du signal vocal dans une catégorie qui contient des formants du signal vocal, et les moyens de comparaison comparent l'amplitude du formant qui contient la catégorie du signal vocal à une fréquence, à l'amplitude du bruit à cette fréquence.

5

16. Procédé selon l'une quelconque des revendications 14 ou 15, dans lequel les moyens (11) qui permettent de déterminer si le signal vocal qui doit être délivré serait intelligible comprennent des moyens qui permettent de comparer l'amplitude du bruit à sensiblement la même fréquence que celle d'un formant du signal vocal, à l'amplitude du formant. 10
17. Système selon l'une quelconque des revendications 14 à 16, dans lequel le signal vocal est représenté par des données qui comprennent des données de paire de raies spectrales (ou LSP, acronyme de Line Spectral Pair), et les moyens (10) qui permettent de modifier les caractéristiques du signal vocal qui doit être délivré par le système de communication vocal comprennent des moyens qui permettent de modifier les données de paire de raies spectrales (ou LSP, acronyme de Line Spectral Pair) qui représentant le signal vocal. 15 20 25
18. Système selon l'une quelconque des revendications 14 à 17, dans lequel les moyens (10) qui permettent de modifier les caractéristiques du signal vocal qui doit être délivré par le système de communication vocal comprennent des moyens qui permettent de modifier la fréquence d'une composante du spectre du signal vocal. 30
19. Système selon la revendication 18, dans lequel les moyens (10) qui permettent de modifier les caractéristiques du signal vocal qui doit être délivré par le système de communication vocal comprennent des moyens qui permettent de modifier la fréquence d'un formant du signal vocal afin de déplacer le formant à une fréquence où l'amplitude du bruit est plus faible. 35 40

45

50

55

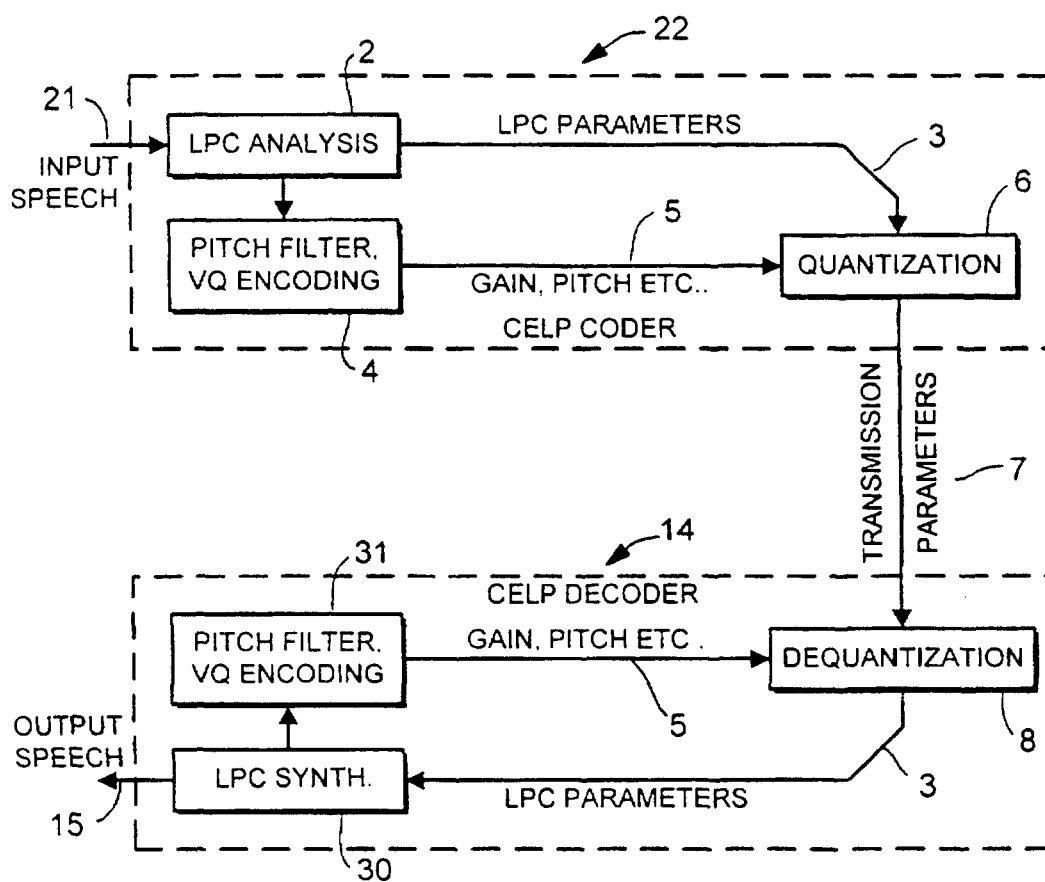


FIG. 1

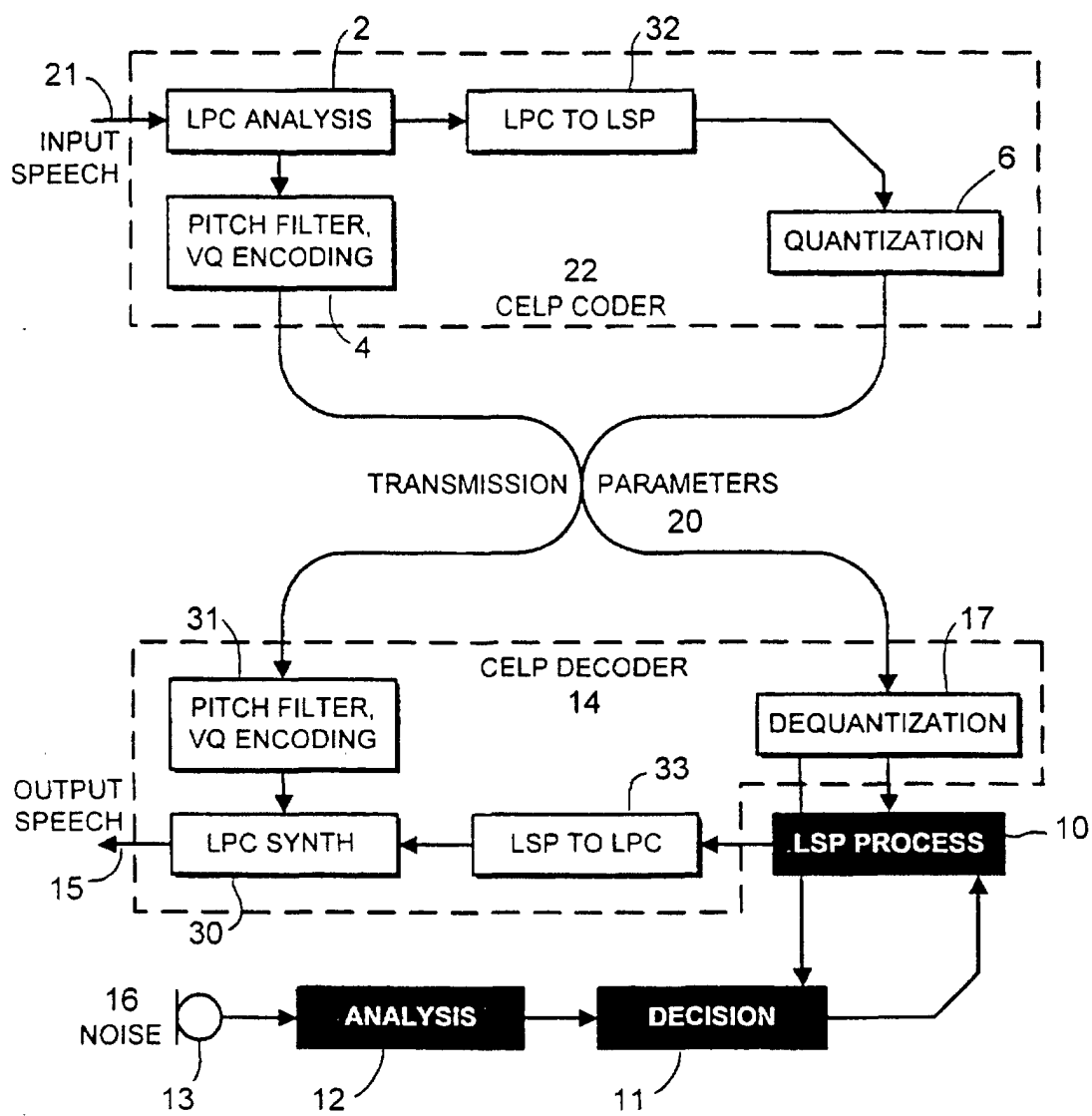


FIG. 2

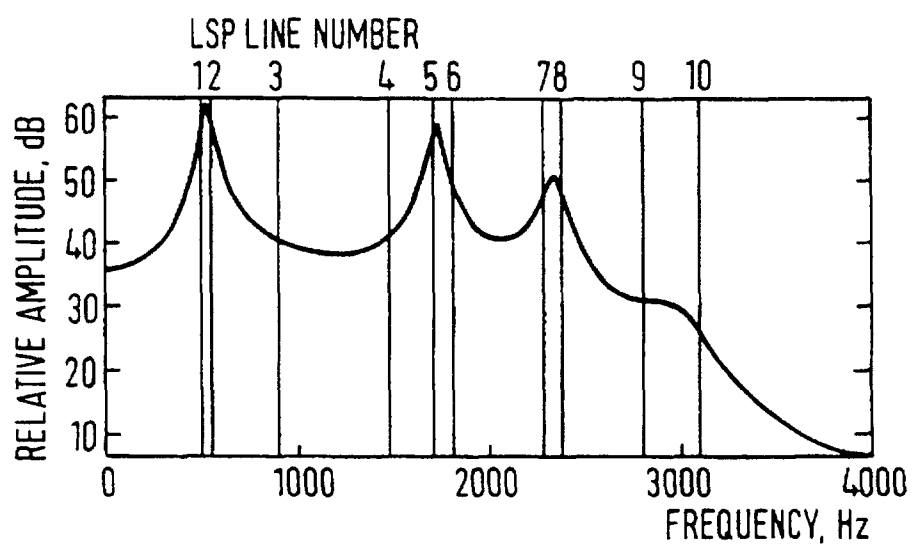


FIG. 3

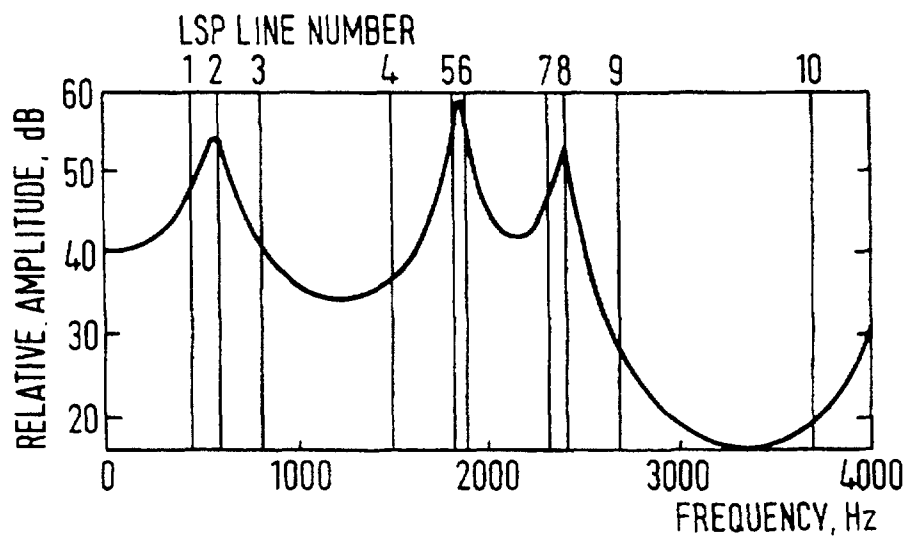


FIG. 4