

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
11 January 2001 (11.01.2001)

PCT

(10) International Publication Number  
**WO 01/03112 A1**

(51) International Patent Classification<sup>7</sup>: **G10L 15/00**,  
11/00, G09B 19/04, 19/06

(74) Agent: **PIZZEYS PATENT & TRADEMARK ATTOR-**  
NEYS; P.O. Box 291, Woden, ACT 2606 (AU).

(21) International Application Number: **PCT/AU00/00817**

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, 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, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(22) International Filing Date: **6 July 2000 (06.07.2000)**

(25) Filing Language: **English**

(26) Publication Language: **English**

(30) Priority Data:  
PQ 1459 **6 July 1999 (06.07.1999)** **AU**  
PQ 3549 **19 October 1999 (19.10.1999)** **AU**

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicant and

(72) Inventor: **QUEST, James [AU/AU]; 1/134 Bandjalong**  
**Crescent, Aranda, ACT 2614 (AU).**

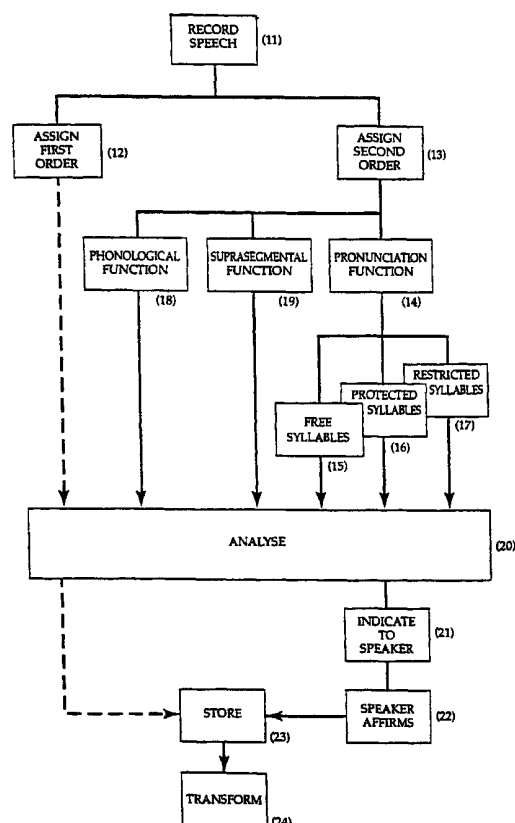
**Published:**

— *With international search report.*

*[Continued on next page]*

(54) Title: **SPEECH RECOGNITION SYSTEM AND METHOD**

(57) Abstract: A method is disclosed of recognising Spoken English (SE) consisting of words having syllables and phonemes, in which the method includes assigning to SE words a first order of signification (12) which includes standardised indicators having agreed meanings independent of the speaker. It also includes designating the syllables in words as being protected, restricted or free syllables which assigns a potential variability to each syllable that may be obtained in connected speech. A second order of signification is also assigned to SE words (13), this second order having variable indicators which have meanings which are generated by the speaker and are dependent on the context of the word in the flow of connected speech.



WO 01/03112 A1



---

*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.*

**"SPEECH RECOGNITION SYSTEM AND METHOD"**

**Technical Field**

This invention relates to speech recognition systems and methods.

5 As used herein the expression "speech recognition" is to be given a broad meaning. The expression "recognition" for example is to be understood to include, inter alia, analysis (including forensic analysis), synthesis and interpretation, coding and de-coding. "Speech" is to be understood to include all speech and voice whether human or artificial. Common expressions used in the art and which are to be regarded as being included within the meaning of the expression "speech recognition" as used herein include voice recognition, language  
10 processing, speech research, spoken-language processing, speech perception, speech synthesis etc.

The invention has particular but not exclusive application to the recognition, analysis and interpretation of spoken English (SE). The invention has more particular application to a range of specific applications including voice recognition technology (VRT) and the teaching and analysis of SE, both as a first language and as a second or other language, hereafter called "English as a second language" (ESL).  
15

**Background of Invention**

Traditional scholarship divides the phonological systems comprising the spoken languages of the world into two main categories - stress timed language systems such as spoken German, Russian, Arabic and Greek, and syllable-timed language systems such as spoken Chinese - both Mandarin and Cantonese, spoken Korean,  
20 spoken Vietnamese, spoken Thai and the Romance languages etc.

The traditionally accepted methods and theory for analysing languages are based on Saussure's concept of the universal linguistic or language sign (hereinafter referred to simply as language sign) and Applied Linguistic's general definition that language is a communication system composed of arbitrary symbols which possess an agreed-upon significance within a community, are independent of immediate context and are  
25 connected in rule-governing ways.

Voice recognition techniques and technology (VRT) are well known and with the power of modern computing have advanced significantly in recent times. VRT has many significant applications including speech to text conversion for sound activated word processing, natural speech synthesis for messages (the so-called Talking Timetable), voice activated hands-free control such as for example can be used for vehicular and appliance  
30 control, voice activated control of bio-medical devices for the disabled, name dialling for telephones etc.

However even with modern computing systems, the development of VRT has lagged its potential, particularly in relation to SE.

It is believed by the inventor that VRT associated with SE has not achieved its potential and advanced to the extent that technology would otherwise permit because modern SE is traditionally regarded as a stress timed language and is analysed for VRT purposes in accordance with the classical Saussurean universal language sign and the Applied Linguistic general definition and other theories of language that are based upon them.  
35

This present invention is based on the inventor's understanding that rather than being a stress timed language, SE has an analytic (hereafter called analytical) phonology in which words and language have two values or orders of signification.

40 The first value relating to words is a standardised or fixed phonetic value and entity which is set by convention, such for example, as defined by the phonetic entries of individual words found in dictionaries using the International Phonetic Alphabet (IPA) and other systems of phonetic notation such as those used by American dictionaries, or as defined by common usage. The word's first value possesses an agreed upon meaning.

In connected SE, the language also possesses standardised indicators that signify readily understood  
45 grammatical, syntactical and linguistic forms, functions and structures (hereafter called 'standardised forms,

functions and structures' or 'readily understood forms, functions and structures' or 'standardised indicators'). These standardised forms, functions and structures exist independently of the speaker and include such things as for example: word order, syntactical formulas, phrases, clauses, sentences, parts of speech, reference words, verb tense, case, voice, aspect, mood, marking the sequencing and relationships of clauses, discourse marking and structuring, time and place markers, register, and socio-linguistic practices, etc. Such standardised forms, functions and structures are established by convention and common usage and may enable more complex social communication to occur within that community of SE speakers.

The forms, structures and functions described above can effectively be regarded as a first value or order of signification relating to language.

The second value relating to words is a variable phonetic value and identity. The second value is variable in time and sound qualities and composition and is defined by the individual speaker at the moment of utterance. The word's second value may possess a further variable meaning, or meanings, that emanates from its first value but which is also relative to the circumstances of the word's immediate "real life" context and the word's place within the flow of connected speech. The variable sound and time imagery obtained by the second value of words and language may also signify some of the various standardised forms, functions and structures of speech, as outlined above.

#### Summary of Invention

The present invention aims to provide an alternative to known speech recognition systems and methods.

This invention in one aspect resides broadly in a method of recognising speech consisting of words having syllables and phonemes, the method including:-

assigning first and second orders of signification to a word;

wherein the first order of signification includes standardised indicators having agreed meanings independent of the speaker, and the second order of signification includes variable indicators having meanings which are generated by the speaker and are dependent on the context of the word in the flow of connected speech.

It is preferred that the speech is spoken English.

In a preferred embodiment the method includes:-

assigning first and second orders of signification to language;

wherein the first order of signification assigned to language includes forms, functions and structures independent of the speaker, and the second order of signification assigned to language includes variable forms, functions and structures which are generated by the speaker and are dependent on the context of the word(s) and/or utterance(s) in the flow of connected speech.

It is also preferred that the method includes:-

analysing the word/s and speech in accordance with the first and second orders of signification.

It is preferred that the variable indicators include the pronunciation of phonemes, syllables and words in the speech. The syllables are preferably categorised as being either free syllables, protected syllables or restricted syllables.

As used herein the expression "protected syllable" means the syllable that customarily carries the main stress in a polysyllabic word but in connected speech may assume any value of stress so long as it is pronounced distinctly. The expression "free syllable" means the syllable that customarily carries main secondary or tertiary stress in a polysyllabic word and also means all monosyllabic words. In words in connected speech the free syllable may assume any value of stress, however prominent or reduced. The expression "restricted syllable" refers to the syllable in a polysyllabic word that does not carry main stress, or main secondary or tertiary stress. In

words in connected speech the restricted syllable may assume a value of stress equal to or less than the free and protected syllables in that word.

It is also preferred that the variable indicators include features of speech such as variations in pitch, tone, harmonic content, volume, duration, rhythm, tempo and the rate of syllables spoken per second.

5 It is also preferred that the variable indicators include other suprasegmental or prosodic features of speech such as variations in the speed of delivery, variations in enunciation, variations in pausing, variations in phrasing and variations in word linking.

It is also preferred that the variable indicators include those grammatical, linguistic and syntactical forms, functions and structures (hereafter called 'variable forms, functions and structures' or 'variable indicators') which are communicated by the speaker by way of variable sound imagery in speech.

10 It is further preferred that the variable indicators include the facts of the immediate context pertaining to the words and utterances in the flow of connected speech.

In a preferred embodiment of the invention the method further includes:-

recording speech spoken by a speaker;

15 indicating to the speaker the meanings of the variable indicators of the recorded speech, and designating or affirming the meanings of the variable indicators indicated to the speaker.

It is preferred that the method also includes storing data representative of analysed words for which the meanings of the variable indicators have been designated or affirmed.

20 In another aspect this invention resides broadly in a method of recognising speech consisting of words having syllables and phonemes, the method including:-

assigning first and second orders of signification to words and language;

25 wherein the first order of signification includes words and language having standardised indicators having agreed meanings and forms, functions and structures independent of the speaker, and the second order of signification includes words and language possessing variable indicators signifying variable meanings and forms, functions and structures which are generated by the speaker and which are dependent on the context of the words and utterance/s in the flow of connected speech. The second order of signification is communicated by way of variable sound imagery in speech.

In another aspect this invention resides broadly in a system for recognising speech consisting of words having syllables and phonemes, the system including:-

30 recording means for recording speech spoken by a speaker;

means for assigning a first order of signification to a word, the first order of signification including standardised indicators having agreed meanings independent of the speaker;

35 means for assigning a second order of signification to a word, the second order of signification including variable indicators having meanings which are generated by the speaker and are dependent on the context of the word in the flow of connected speech;

indicating means for indicating to a speaker the meanings of the variable indicators of the recorded speech, and

designation means whereby a speaker designates or affirms the meanings of the variable indicators indicated by the indicating means.

40 In another aspect this invention resides broadly in a system for recognising speech consisting of words having syllables and phonemes, the system including:-

recording means for recording speech spoken by a speaker;

45 means for assigning a first order of signification to words and language, the first order of signification including standardised indicators having agreed meanings, and standardised forms, functions and structures independent of the speaker;

means for assigning a second order of signification to words and language, the second order of signification including variable indicators possessing meanings, and forms, functions and structures which are generated by the speaker and are dependent on the context of the word in the flow of connected speech;

5 indicating means for indicating to a speaker the meanings, and forms, functions and structures of the variable indicators of the recorded speech, and

designation means whereby a speaker designates or affirms the meanings, and forms, functions and structures of the variable indicators indicated by the indicating means.

It is preferred that the system also includes analysing means for analysing the words and speech in accordance with the first and second orders of signification.

10 It is further preferred that the system includes storage means for storing data representative of analysed words for which the meanings, and forms, functions and structures of the variable indicators have been designated or affirmed.

In a further aspect this invention resides broadly in a method of teaching how to speak a language, the method including:-

15 assigning first and second orders of signification to a word, the word having syllables and phonemes, the first order of signification including standardised indicators having agreed meanings independent of the speaker and the second order of signification including variable indicators having meanings which are generated by the speaker and are dependent on the context of the word(s) in the flow of connected speech, and

practicing speaking using different variable indicators in the second order of signification.

20 It is preferred that the method includes:-

assigning first and second orders of signification to language, the first order of signification including forms, functions and structures independent of the speaker and the second order of signification including forms, functions and structures which are generated by the speaker and are dependent on the context of the word(s) in the flow of connected speech.

25 It is preferred that the method also includes practising speaking in order to acquire the preferred respiratory, cognitive and vocal skills, and the preferred skills of physical and mental co-ordination, in using variable indicators in the second order of signification.

The method may further include analysing connected speech that exhibits different variable indicators in the second order of signification for the purposes of recognising and evaluating the speech for more complete meaning.

30 It is preferred that the language is spoken English and that it is taught as a first language.

It is also preferred that the language is spoken English and that it is taught as a second language.

#### Description of Drawings

35 In order that this invention may be more easily understood and put into practical effect, reference will now be made to the accompanying drawings which illustrate a preferred embodiment of the invention, wherein:-

FIGS 1A to 1F diagrammatically illustrate some variable word stressings of the word "disappointing" and illustrate free, restricted and protected syllables in connected speech;

40 FIGS 2A to 2G diagrammatically illustrate seven commonly understood uses of pitch and tone in SE, providing various examples of their use and readily understood meanings;

FIGS 2H to 2J diagrammatically illustrate the uses of variable pitch and tone in spoken phrases to compress additional meaning into speech in the second order of signification;

FIG 3 is a flow chart of a method of recognising speech in accordance with the invention, and

45 FIG 4 is a schematic block diagram illustrating a system for recognising speech in accordance with the invention.

### Description of Preferred Embodiments of Invention

Before providing a more detailed description of the preferred embodiments of the methods and description of this invention, a description of the inventor's understanding of underlying principles will be given, first at a more general level and then in summary.

#### - Synthetic and Analytical languages

The key languages that helped to shape modern English are what linguists call synthetic languages. Modern German, for example, is a synthetic tongue, as was Latin.

A synthetic language is one that possesses a complicated system of grammar dependent upon the use of inflections. Inflections are word endings or affixes that denote things such as gender (inanimate objects having a sex), case, voice, verb tense and number. In spoken German, the more complex an idea to be expressed, the more inflections are required thereby producing words of many syllables. The modern Germanic tongues thus continue to use inflections as an integral part of their grammatical systems.

By contrast, Modern English is called a syncretic or analytic language (hereafter referred to as an analytical language or an analytical system).

An analytical language is oppositional to the synthetic languages in its core philosophy and drive. A key feature of an analytical system is that it does not depend upon inflections for expressing complex ideas and meanings. The system of inflections was largely abandoned sometime during the period currently referred to as Middle English. Instead an analytical language follows the reductionist path in seeking to express more complex ideas and meanings with shorter words, through the use of simpler grammatical forms and structures, through the flexibility of the function of words that may assume various parts of speech, (a flexibility which also enables the constant creation of catch phrases, context-specific jargon, and new descriptive words, phrases, terms and expressions) and through expedient pronunciation habits and practices such as word linking..

In the spoken word of an analytical language such as modern SE, the system of fixed word order (e.g, subject + verb + object: "He loves Mary", or subject + verb + complement: "She feels sick") creates simple grammatical structures and readily understood syntactical formulas which can support complex meanings.

This hitherto insignificant distinction between synthetic and analytic, in the first instance, places SE outside the family of synthetic Germanic tongues that are governed by a stress timed phonology. The hallmark of the entire English language system, both written and spoken, is that it is an analytical system. Ergo, the connected speech of everyday English is governed by an analytical phonological system.

#### - Stress timed languages

Stress timed language systems include spoken German, Russian, Arabic and Greek. (SE has also traditionally been regarded, incorrectly in the view of the inventor, as a stress timed language). These spoken language systems are governed by a different timing principle than syllable timed phonologies.

The main tenets of a stress timed language are that:

(i) The function of main stress is to give prominence to certain words, or even one word, within an utterance, main stress falling on the salient syllable of the stressed word.

(ii) Main stress beats fall at roughly regular intervals of time within connected speech with weaker stress beats falling on the words and syllables in between. (In English Linguistics, the chunk of speech between two main stress beats is often called a foot, with feet said to be "about the same".)

(iii) Stress, therefore, produces the one clear, underlying rhythm in streams of natural speech as it is organised around a stable timing principle.

(iv) Stress and rhythm, in their own right, are incapable of, or severely limited in, expressing contrasts in meaning, and

(v) The phonological system imposes such order from within.

## 5                    -                    **The Analytical Phonology**

Because modern SE is not a stress timed language as are those Germanic tongues governed by a stress timed phonology, it is beneficial to consider SE as being represented by a new category of phonological system - the analytical phonology. Another example of an analytical language system is written Chinese which uses ideograms and not words composed of characters from an alphabet.

10                    The new category of the Analytical Phonology redefines the nature, general properties and general principles of the English language phonological system that apply where English is spoken as the sole mother tongue of the majority, ie Australia, Canada (other than Quebec), New Zealand and the United States of America - countries which were colonised and populated by England - and in Great Britain and Ireland.

15                    The primary principle and purpose of an analytical language system, in either written and spoken language, are its indomitable urge to express increasingly more complex meanings and ideas while at the same time reducing the actual number, form and length of words as well as reducing and simplifying the standardised forms, functions and structures of speech required to communicate these meanings. It is in modern SE that the analytical principles and purpose of the language are at their most potent.

20                    Put simply, the analytical phonology of modern SE constantly seeks ways and means to say more with less.

## -                    **Stress in Spoken English**

25                    The traditional classification of English as a stress timed language is believed by the inventor to be flawed because connected English speech is not governed by any of the principles of stress timing. This is because stress, or more accurately, vocal prominence, in SE is fixed to no one stable timing principle but is deregulated and highly unpredictable in its distribution and its timing formations in connected speech. Stress, or prominence, in SE follows no one phonological principle of timing and is capable of adopting either syllable or stress timed patterns and formations, or any other irregular patterns or formations obtainable in speech.

30                    On the definition of 'stress' it need be noted that 'stress', as currently understood, is largely an abstract linguistic concept. For our purposes here 'stress' shall be taken to mean those prosodic features of speech that the speaker may use to give vocal prominence to sounds, phonemes, syllables, words and utterances in connected speech. Specific prosodic features that can combine to constitute 'stress' include, for example, volume, pitch, tone, harmonic content, timing, rhythm, clarity, and the duration of a sound, phoneme, syllable, word or utterance in order that it may stand out and gain prominence. Wherever the term 'stress' is used within this document this  
35                    meaning is to be implied.

Consequently, within the definition and parameters of the free syllable, the protected syllable and the restricted syllable, the gradations of stress in speech are otherwise freely transferable between phonemes (particularly the vowel sounds), syllables and words within utterances.

40                    There are many and various gradations of stress obtainable in natural connected speech, not merely weak and strong. In general, stress in SE cannot be regulated nor categorised according to any fixed measure of gradation that the phonological system imposes.

The gradation and composition of stress that may attend sounds, phonemes, syllables, words and phrases are relative to the utterance, having due regard for the immediate context of the utterance within its particular stream of connected speech. It is on this basis that the listener, amongst other considerations,



evaluates the gradations, composition and contrasts of stress audible within the speaker's speech and judges what is being made prominent by the speaker in terms of both sound and words.

Stress, as defined above, is capable of signifying meaning or facilitating it.

5 The speaker may use stress to make certain sounds, syllables, words and phrases prominent in sound and meaning. While stressed words and speech may, potentially, achieve vocal prominence, main stress is not the only method by which words may stand out in speech. Sometimes 'unstressed' or understated words and speech may also signify prominence if weakly stressed speech can stand out in the stream of spoken discourse (for example, a weakly stressed sound, syllable, word, or words, embedded within a passage of highly stressed speech can achieve prominence). Prominence may also be achieved by way of the other phonological and  
10 prosodic features of everyday connected speech. In making prominent certain sounds, phonemes, syllables, words and phrases in speech the speaker may wish, for example, to highlight certain words, place special emphasis and focus on certain words and parts of speech, or juxtapose, rank, contrast, infer, counterpoint or complement meanings and ideas embodied within and between utterances.

15 The uses to which stress in everyday SE may be put are at the behest of the individual speaker, precisely for the purposes of generating and producing further meaning, and/or for the purposes of reducing, substituting, or simplifying the standardised forms, functions and structures of speech.

As a consequence of the nature of stress in everyday SE, the rhythmic organisation and tempo of connected speech are also highly unpredictable.

Therefore, stress in natural SE may be termed variable.

20 Because stress in SE is variable, the phonology of SE is analytical, and vice-versa, making it a language system markedly different from both the syllable timed and stress timed languages. SE is thus appropriately categorised independently as representative of the analytical phonology.

#### - Variable Stress

25 The basic phonological functions and suprasegmental features of connected English speech are such that the fundamental materials of variable sound and time with which the native English speaker may construct speech, are quite rich and manifold. These variable sound and timing variations in connected English speech follow from the inescapable fact that stress in the spoken English language is variable.

30 The basis of variable stress and a 'variably timed' phonological system is rational and logical and is the consequence of the spoken language system being governed by no one fixed principle of timing.

Variable stress can be conceptualised as a phonological freedom rather than an imposition or restriction, and which is granted rather than imposed from within.

35 This freedom becomes manifest and widely exploited in everyday speech. Variable stress and its consequences on the phonological and suprasegmental features of everyday speech allows individual speakers enormous freedom to engineer the sound qualities and patterns of speech enabling them to signify meanings and ideas in tremendously creative, idiosyncratic and inventive ways. It also enables some of the standardised forms, functions and structures to be communicated by way of variable sound and timing imagery rather than by words and verbal formulations.

#### 40 - Sound and Timing Variables in Spoken English

Amongst the key phonological functions of SE dependent on variable stress (hereafter called the key phonological functions or the key phonological features) include variable timing and duration, variable volume, variable rhythm and tempo and variable pitch and variable harmonic content. Other suprasegmental and prosodic features of speech effected by variable stress (hereafter called other suprasegmental features) include variable  
45 speed of delivery, variable word linking, variable enunciation, variable pausing and variable phrasing.

All of the above key phonological features and phonological functions and the other suprasegmental features of everyday connected English speech are inextricably linked to variable stress. 'Stress' is well understood to have a commanding role to perform in the organising functions of the phonological system. However because stress in SE is variable, the interconnected phonological functions are also variable.

- 5           Because stress is variable the role of the other suprasegmental features of speech becomes important. These features generally are variable and/or optional. Their effects on the phonological functions in connected speech greatly heighten the system's overall capacity and scope to accommodate greater sound and timing variations and contrasts which variable stress enables.

- 10           All of the sound functions and features of everyday speech are capable of variation within a wide and legitimate parameter of sound & timing variability that the English phonology permits and encourages and which, crucially, the English language phonetic system tolerates and accommodates.

The factors of sound and timing variability in connected speech work at the basic level of the phoneme, the syllable and the word and at the more general level of spoken phrases and connected speech.

- 15           Sound and timing variations effect or change the way the native speaker pronounces phonemes, syllables, words and phrases in ways the listener can hear.

The listener is constantly keeping track of the particular audible qualities of each phoneme, syllable, word or phrase, and on a comparative scale of values, is registering and measuring variables such as:

- How strong, stressed or prominent? How weak? is each phoneme, syllable, word or phrase when compared with another? The same or different?
- 20           (Words + variable stress)
- How loud? How soft? is each phoneme, syllable, word or phrase? The same or different?
- (Words + variable stress + variable volume)
- How long? How short? is each phoneme, syllable, word or phrase? The same or different?
- (Words + variable stress + variable duration and timing)
- 25           How fast? How slow? is each phoneme, syllable, word or phrase? The same or different?
- (Words + variable stress + variable speed of delivery, duration and timing)
- How high? How low? is each phoneme, syllable, word or phrase? The same or different?
- (Words + variable stress + variable pitch)
- How wide and deep are tone groups? How narrow or shallow are tone groups around each phoneme,
- 30           syllable, word or phrase. The same or different?
- (Words + variable stress + variable tone and variable harmonic content)

Also, the listener follows the flow of speech in order to gauge other sound features that can produce variation, paying attention to things such as:

- Are words linked? Are words not linked?
- 35           · Are phonemes, syllables, words and phrases distinct? Are they indistinct?
- Is any particular sound and timing variation repeating itself?
- Is any particular sound and timing variation creating prominence in some way?
- What are rhythmic and tempo changes signifying?
- What are pauses signifying, framing or announcing?
- 40           · Are sound and timing variations signifying any forms, functions and structures of speech?

- 45           A fundamental axiom of the analytical phonology is that the native speaker is free to apply any of the many variable phonological functions and suprasegmental features in the pronunciation of phonemes, syllables and words at choice which are obtainable in speech for the purposes of creating further meaning provided that such variations do not negate the agreed-upon meanings of the words nor nullify the standardised and readily understood forms, functions, and structures of speech..

To these ends and within this fundamental rule, further meaning can be generated when the phonemes and syllables that constitute the word - still an arbitrary symbol possessing a static agreed-upon meaning - are varied in sound and timing according to the speaker's pronunciation and in ways the listener hears and registers as meaningful.

5 Having regard to the speaker's manner and habits of pronunciation, the variation of sound and timing and the application of the other variable and optional suprasegmental features at the level of the phoneme, syllable and word, cannot but effect the structure, qualities and organisation of spoken phrases and larger passages of discourse as connected speech progresses. As noted, sound and timing variations may also signify various forms, functions and structures of discourse.

10 Within phonemes, syllables, words and phrases, and within the general flow of connected speech, specific sound and timing variations (hereafter called variations) will combine to create discernible sound and timing contrasts (hereafter called contrasts). Therefore, in terms of the sound imagery of SE, the following equation applies:-

15 Phonemes, syllables and/or words + X number of variations =  
Words and phrases + Y number of appreciable contrasts in speech.

The potential number of variations and specific combinations of variations capable of creating discernible contrasts within connected English speech is untold. This enables the same English word to obtain a multitude of different sound images in speech, either as a stand alone word or as it finds itself placed within the traffic of connected speech. This is simply because with so many variations, choices and options available the potential combinations of variations able to produce appreciable contrasts are innumerable.

20 Nonetheless, in monitoring and comparing the qualities and composition of specific phonemes, syllables, words and phrases and in monitoring the overall changes and variations within connected speech (as outlined above) the fluent listener can recognise contrasts and from this distinguish, within the context of the conversation, what is being made salient in terms of meaning by the speaker.

25 On the level of the phrase, the three word utterance "I love you" consisting of three monosyllabic words, can potentially obtain virtually unlimited variable sound images in speech according to such factors as the phrase's internal contrasts of variable stress, variable pitch, variable tone, variable timing and duration, variable volume, variable speed of delivery, the variable degrees of distinctness and clarity obtainable, how the words might or might not be emphasised, if they are linked or unlinked to each other and the surrounding utterances, and the duration and qualities of the pauses that delineate the words and phrase.

30 However irrespective of what legitimate variations and contrasts may attend the phrase, the actual agreed-upon meanings of the words "I love you" remain constant and the basic structure of the utterance (in this case, subject + verb + object) remains intact.

35 Therefore, the various renderings of the phrase "I love you" can generate extra meaning according to the legitimate variations and contrasts that attend each utterance without negating the word's agreed-upon meaning and without nullifying the standardised forms, functions and structures of the phrase. In this instance -- as in all cases where additional meanings and various forms, functions and structures (such as register, or the implied relationship/s between the subject, the verb and the object of the utterance) are signified by variable sound imagery -- the specific meanings, forms, functions and structures generated by sound and timing variations and contrasts can only be properly understood within the frame of reference of the utterance's immediate context. This is because such meanings are dependent on and relative to both the form, function and structure of the utterance, the lexical content of the utterance, and the unique facts that pertain to the circumstances of the utterance's immediate context.

40  
45

Even more legitimate variations and contrasts can be obtained when a particular phoneme, syllable, word or phrase is compared to other utterances, subject to the same factors of variation and contrast, whether these other utterances are within or without the immediate context.

5 These further meanings and implied forms, functions and structures, which may be termed second order meaning, are communicated to the listener at the same time that the word's agreed-upon meaning and the standardised forms, functions and structure/s of the utterance are signified.

The agreed-upon meaning of the actual words and the standardised forms, functions and structure of the utterance which remain constant throughout may be referred to as the first order of meaning.

10 The legitimate variations and contrasts, or variable sound imagery, that can produce second order meanings not only create further meaning/s in the mind of the listener/s but may also create new meaning/s in the mind of the listener/s

15 Since the particular combination of sound and timing variations that might create a particular contrast, or cluster of contrasts, cannot be pre-defined - as this is something that is initiated and articulated by the individual speaker in the moment of utterance and within the extemporaneous circumstances of the particular context - we are unable to predict every possible context nor classify the mood, manner of speech and temper of mind of every individual speaker.

20 The onus then falls upon the listener to be able to recognise, register, decode and interpret variations and contrasts for second order meanings. This is something at which the fluent native speaker is proficient. In short, sound and timing variations and contrasts in the sound imagery, or soundscape, of everyday connected English language speech work productively to be the signifiers of second order meanings.

This system through which the production of second order meaning in SE occurs may be termed the second order of signification.

25 The traditional and orthodox linguistic system of producing sounds that signify static agreed-upon meaning and standardised forms, functions and structures, common to all languages, and as conceptualised by Saussure, may be termed the first order of signification.

#### - Analytical Phonology and the Phonetic System

30 The analytical nature of the phonology of SE cannot but effect its phonetic system because the language depends upon its phonological system and its phonetic system co-operating in order that the second order of signification may function properly. Moreover, the variable phonological functions and suprasegmental features of everyday SE operate at the level of the phoneme.

The phonetic system of SE differs markedly from the phonetic systems of the non-analytical language systems in numerous ways including the following:-

35 The way words are spelt in English bears no logical correlation, nor readily understood systematic connection, to the way the words are pronounced. The phonetic pronunciation of words is set by common usage and not by the word's spelling. This creates a fundamental dislocation or disjuncture between the written and spoken systems of the English language at the level of the spellings of words and their pronunciations with the onus falling heavily on the practices and conventions of common usage to determine the acceptable and comprehensible pronunciation of words with these practices and conventions often being in conflict and flux.

40 The phonetic system of SE has the ability to tolerate marked and sometimes radical sound and timing variations within individual phonemes which having no one fixed principle of timing the system can permit. The phonetic system will indifferently accept legitimate variations without negating the static agreed-upon meaning of the word and without rendering the individually varied phoneme that constitutes the word incomprehensible to the listener.

45

This remarkable tolerance of the English language phonetic system is exemplified by the ability of all of the six written vowels **a, e, i, o, u** and **y** (when a vowel sound) to assume or be influenced by the reduced vowel sound /ə/ in certain and numerous syllables and words and in liberal application, such as the short vowel sounds in free and restricted syllables.

5           The ability of the system to tolerate hybrid sounds and phonemes in standard words and blurred sounds, phonemes, glides etc., between linked words, and the many unpredictable phonemic effects that come when syllables and words are made prominent or are linked. Many such sounds cannot necessarily be placed in the standard IPA chart but nonetheless are perfectly understandable to the listener when they occur in words in a context in the flow of speech.

10           It also enables any of the 44 sounds of SE, as defined by the International Phonetic Alphabet (IPA), to be stretched or elongated according to any timing format or aesthetic purpose that the individual speaker can successfully obtain in natural speech. This function particularly effects the vowel sounds of natural SE which are the phonemes most commonly varied.

15           The absence of any one fixed timing principle enables ordinary vowel sounds to become diphthongs, and diphthongs to become "triphthongs" and so on. The ability to elongate sounds, vowels in particular, is restricted only by the respiratory limits of the speaker and defined by the context of the utterance and according to the speaker's intention to create second order meaning of some kind.

20           Unlike the phonetic system of other language systems, the phonetic system of SE does not have a basic, fixed and limited stock of phonemes that are used to construct syllables and words. Instead the phonetic identity and the quantity of legitimately varied phonemes actually available for the purposes of constructing intelligible syllables and words in everyday SE are so numerous they cannot be reliably counted.

25           It is the belief of the inventor that the fundamental disjuncture between the spellings of words and their pronunciation accounts, in part, for the necessary linguistic pre-conditions that have enabled the stock of phonetic sounds available for the construction of intelligible English words to expand and become "un-delimited".

In practice this means words, as phonetically defined, have two values:

- 30           (i) a standardised or fixed phonetic value and identity set by convention and as defined by the phonetic entries of individual words found in dictionaries using the IPA, and other systems of phonetic notation, and
- (ii) a variable phonetic value and identity, in terms of the phoneme's time and sound qualities and composition, something that is defined by the individual speaker in the moment of utterance and which is relative to its standardised phonetic value as well as to the circumstances of the word's immediate "real life" context and its place within the flow of connected speech.

35           The standardised phonetic value of a particular phoneme is the necessary reference point by which the phoneme's second variable value may be recognised, evaluated and interpreted for the purposes of producing second order meanings. When the standardised phonetic content of sounds, syllables or word/s is appreciably varied then the speaker may well be making that sound, syllable or word/s 'prominent' in speech in order to signal second order meaning, or meanings, to the listener.

40

#### - **Analytical Phonology and Word Stress**

The analytical nature of the phonology of SE also effects word stress. This is because the variable phonological functions and suprasegmental features of everyday SE operate at the level of the syllable and the word. As words enjoy two phonetic values, one being standardised and the other variable, likewise words enjoy two values in the way they are stressed.

45

The first "word stressing" is standardised by convention and defined in the phonetic entries of individual words found in dictionaries which usually mark which syllables within polysyllabic words customarily assume main, secondary and weak stress.

5 The second "word stressing" is a variable that is defined by the individual speaker in the moment of utterance and which is relative to the word's standardised stressing as well as to the circumstances of the word's immediate "real life" context and its place within the flow of connected speech.

The system of variable word stressing, as referred to above, operates within certain limits and parameters. These limits are defined by the three categories of syllables a polysyllabic word may obtain within the flow of everyday connected speech.

10 The three categories of syllables are:-

(i) The distinct or protected syllable.

15 This is the syllable that according to the normal "dictionary" standard carries the main stress of the word. When the protected syllable does not carry main stress, or vocal prominence, within an utterance it may then assume any gradation of stress so long as it is pronounced as distinct as is necessary so as to not to disturb the word's internal distribution of stress to the extent that the word's agreed-upon meaning becomes negated or unclear. Both syllables in compound nouns can usually be regarded as protected syllables.

(ii) The free syllable.

20 This syllable can assume any gradation of stress within an utterance, from being the most reduced sound to the one carrying main vocal prominence and any point in between. It is different from the protected syllable, as it is the syllable which dictionaries define as carrying secondary stress of one level or another. In many cases within the context of the phrase it may carry more stress than the syllable/s that by normal definition should carry the main stress of the word. When this occurs, or in instances when the free syllable is a reduced sound, it is usually because the speaker wishes to communicate a second order meaning of some kind. It need be noted that virtually all monosyllabic words may enjoy any gradation of stress, vocal prominence, or duration of time obtainable within natural connected speech, the small function words (such as a, an, is, the, of, to, in etc) included, for the purposes of producing second order meanings. Therefore monosyllabic words can generally be regarded as free syllables. Inflections in polysyllabic words of more than two syllables can generally be regarded as free syllables not carrying the reduced sound /ə/.

(iii) The restricted syllable.

30 This is the syllable in a word that by normal definition is unable to support main, secondary or tertiary stress. In general, syllables carrying the reduced sound /ə/, and inflections in polysyllabic words of two syllables are restricted syllables. The restricted syllable will very rarely assume the only main stress within a polysyllabic word. When it does assume main stress it is to give the word a variable sound or timing value in order to generate a particular second order meaning for the purposes of, say, mimicry or sarcasm. Within the sound environment of the utterance the stress gradation or vocal prominence of the restricted syllable in a word will normally never be greater than the protected and/or free syllables within that word and is most often of an appreciably weaker gradation. This does not mean, however, that the stress gradation of a particular restricted syllable in one particular word - when appreciated within the sound context of the entire utterance - need necessarily be relatively weaker or stronger than any of the syllables of other words in that same utterance since the phrase may be supporting a multiplicity of second order meanings that the speaker wishes to communicate simultaneously.

40

Brief reference will be made to FIGS 1A to 1F which illustrates the potentially variable word stressings of the word "disappointing" with reference to free, restricted and protected syllables in the analytical phonology.

FIG 1A shows the standardised dictionary stress pattern for the polysyllabic word "disappointing" which consists of the four syllables: **dis**, **ap**, **point** and **ing**.

However the word is subject to much variation in the traffic of connected speech when placed in different phrases carrying different second order meanings.

Thus as illustrated in FIG 1B which shows the potential variation of syllables in connected speech, the first syllable **dis**, which dictionaries define as having secondary stress, is a free syllable. The last syllable, the inflection **ing** in this case, is also a free syllable; the second syllable **ap**, carrying the reduced sound /ə/, is the restricted syllable, and the third syllable **point**, carrying main stress, is the protected syllable.

In the phrase "The result was disappointing" illustrated in FIG 1C, the stressing of the word "disappointing" is as the dictionary would prescribe.

In the phrases "A disappointing result" seen in FIG 1D and "That's disappointing" seen in FIG 1E, the word "disappointing" is still prominent and is stressed variably according to the speaker's construction of the phrase and the particular second order meaning he or she wishes to convey.

In the phrase "Very disappointing indeed" seen in FIG 1E, the word "disappointing" does not obtain prominence in the phrase at all. The speaker chooses to place the vocal emphasis on the words **VERY** and **indeed**. In the examples where the standardised stressing of the word "disappointing" has been varied it is done so in order to generate a second order meaning within the utterance so as to signify, for instance, an emotional or subjective content or meaning.

A word will deviate from its standardised word stressing and assume a variable word stressing and variable gradations of stress and prominence in connected speech for the purposes of signifying second order meaning.

As is the case with phonemic variation, the standardised stressing of a particular word, as the dictionary or common usage would define it, is the necessary reference point by which the word's second variable stressing may be recognised, evaluated and interpreted for the purposes of producing second order meanings. When the standardised stressing of a word is appreciably varied then the speaker may well be making that word 'prominent' in speech in order to signal second order meaning, or meanings, to the listener.

#### - Analytical Phonology and Intonation

By its liberating effect, variable stress in SE enhances the use of pitch and tone in everyday speech which, by extension, is likewise tethered to no one dominating or fixed principle of timing. In general, variable pitch and tone are a means by which the speaker may produce second order meanings in speech through the melodic interpretation of phonemes, syllables, words and phrases.

Variable duration and timing which allow for the reduction of sounds will also permit the elongation of sounds. In creating this extra time in the pronunciation of phonemes, syllables, words and phrases the individual is free to apply appreciable variations and contrasts of tone and harmonic content (hereafter called 'tone' or 'contrasts of tone'), and pitch in speech.

The use of variable pitch and tone in SE is a means by which phonemes, syllables and words may gain prominence in meaning within connected speech even if these phonemes, syllables and words are unattended by main stress, or even secondary stress (that is, unattended by, for example, prominent volume, emphasis or duration). Nor is the use of variable tone and pitch restricted to the end of phrases only.

At the more general level of connected speech the speaker's use of variable tone and pitch adds to the totality of meanings a spoken phrase may simultaneously support.

Pitch and tone, accommodated within any timing pattern obtainable in speech, create legitimate sound and timing variations that will not negate the agreed meaning of words, nor nullify standardised forms, functions and structures, but moreover serve as the signifiers of second order meaning.

In all cases the listener's and observer's knowledge of SE, the speaker/s and the immediate context will enable such second order meanings to be understood.

Also listeners and observers rely on commonly understood indicators of pitch and tone widely used in connected speech as a reference point by which legitimate sound and timing variations and contrasts may be recognised, registered, decoded and evaluated for meaning.

Reference will be made to FIGS 2A to 2G which illustrate some commonly used and readily understood uses of tone and pitch in connected speech. These serve as further standardised indicators of speech.

FIG 2A shows the use of a rising pitch and tone, commonly understood to signify a yes/no question, a clarifying question, a request for repetition, and interested feedback. In general, rising pitch and tone at the end of phrases or words commonly signifies a sense of doubt, incompleteness, or a need to know more on the part of the speaker.

FIG 2B shows the use of a low rising pitch and tone commonly employed by speakers when reading items from a list to signify that the list is not yet complete, and commonly to signify more neutral feedback or mild interest in what is being said.

FIG 2C shows the use of flat or level pitch and tone in speech, commonly employed to signify disinterest, boredom or sarcasm. In general, level tones will also commonly attend routine or impersonal conversational exchanges.

FIGS 2D and 2E show the use of rising-falling and falling-rising pitch and tones, respectively. Both tone patterns are commonly understood to signify greater emotional content and expression attending the speaker's speech, or to signify contrasting or competing meanings, or to signify a change in register, mood or conversation topic on the part of the speaker.

FIG 2F shows the use of the falling pitch and tone in speech, commonly understood to signify completion, such as when reading the final item in a long list. A falling pitch and tone commonly attends information or "wh-" questions (such as What, When, Where etc) which are asked in the expectation that the answer will be readily provided. Falling pitch and tone will also commonly attend declaratives, statements of fact, and mild apologies the speaker is making. In general, a falling pitch and tone will commonly signify completion, and an absence of doubt in regard to the speaker's utterance.

FIG 2G shows the use of a sharp falling pitch and tone which commonly attends stronger apologies, imperatives, firm statements and declaratives. In general, a sharp falling pitch and tone commonly signify finality, certainty, completion and commonly attend utterances that signify there is no doubt at all in the speaker's mind about what is being said.

The use of variable pitch and tone in connected speech by the fluent speaker enables compression of meaning to occur in utterances when the words of the first order of signification communicate one meaning while, at the same time, the appreciable variations and contrasts of pitch and tone within the second order of signification signify other meanings.

Frequently, speakers in posing yes/no questions, which may commonly adopt a rising tone which signifies doubt, may frame their yes/no question in a falling tone, signifying the opposite: certainty.

Reference will be made to FIG 2F which illustrates this point. The question **Did you murder your wife?** combines both a yes/no question, signified by the words and the grammatical form of the utterance - the phrase's first order of meaning - while at the same time, there is an underlying declarative "I'm certain you did" that is being implied by the prominent falling tone of the speaker's speech, the utterance's second order of meaning, which is signifying the speaker's certainty and not doubt as regards the answer to his or her own question.

Declaratives are subject to the same effects of variable tone and pitch.

Reference will be made to FIGS 2F and 2G. Speaker A's utterance: **I'm going to stop drinking**, may express opposing ideas and meanings simultaneously. The words themselves, the phrase's first order of



meaning, express an apparently self-confident resolution, particularly as the words 'going' and 'stop' are stressed. However, a rising tone and pitch at the end of the phrase around the word 'drinking' could betray the speaker's own sense of doubt or commitment to his or her stated intention, particularly noticeable if the phrase is strongly stressed.

- 5 Similarly, in FIG 2G Speaker B's response: **I'm sure you will**, with the contours of intonation curling upwards on the last word pitches a statement of fact on an uncertain note, subtextually suggesting Speaker B's sense of doubt, mistrust or ambivalence concerning Speaker A's utterance.

- 10 Another example of this is the Australian tendency to finish statements on a rising tone (called the 'mid rising pitch') which combines two meanings through both first and second orders of signification. The speaker is telling the listener something, narrating some past event, or relating information, by way of the words and forms, functions and structures of the first order of signification, while the voice is producing second order meaning by way of variable sound imagery: ie, the mid rising pitch and tone at the end of each phrase is embedding the declarative in an habitual yes/no question: signifying a kind of abiding doubt or tentativeness on the speaker's part or a need to constantly check with the listener that he or she understands and is engaged in the conversation.

- 15 A fluent speaker may also express a pejorative idea, a criticism, reproach or complaint by exploiting the opportunities the second order of signification offers. Rather than deliver an insult in actual words, the speaker may choose to signify the pejorative content through his or her voice in order to achieve the same effect without actually saying anything pejorative in actual words.

- 20 For instance, the speaker may articulate by way of the first order of signification the utterance **You look fabulous!** which would seem a plain and simple compliment. But were the pitch and tone variations to combine with a weak stress beat to create variable sound imagery exhibiting a tepid, or prominently understated, mid-falling or flat tone around the prominent syllable in fabulous this could give the phrase a sarcastic meaning. Here, second order meaning, expressed by the voice, is signifying: **You (don't) look fabulous** or **You look horrible**.

- 25 The fluent speaker of SE is well acquainted with the everyday practices of using pitch and tone in this way. Within the wide range of variable sound imagery of SE available for the purposes of producing second order meanings, speakers are frequently presented with an utterance that economically compresses multiple ideas which express meanings that realise no literal form. Phrases signifying moods, meanings and ideas of, for example, disappointment, anger, annoyance, frustration, pleasure, desire, sarcasm, contempt and so on, may easily be expressed phonologically in tandem with the static agreed meanings and the standardised forms, functions and structures of the utterances' first order of signification. Variable pitch, tone, duration, timing and stress may also be used as signifiers of second order meaning by speakers within certain regional sub-varieties of SE and within certain idiolects of SE as a kind of 'in-group' speech code for the purposes of identifying one member of a certain speech community or 'in-group' with his or her peers. This is particularly prevalent among younger speakers of SE, such as second generation migrant youth in Sydney, the so-called 'Valley Girls' of Los Angeles and young followers of the Australian TV soap opera 'Neighbours' in the United Kingdom, who have adopted the Australian tendency of the 'mid rising pitch' in natural speech, hitherto unknown in that part of the world.

- 35 Thus, subtextual content, emotion, irony, idiomatic codes, complementary moods and meanings can be compressed within the one spoken English phrase without need of the speaker formulating a new spoken phrase in order to express these further meanings. In extended discourse various forms, functions, and structures may also be signified by way of the variable sound imagery of second order meaning.

In this way the analytical purpose of the spoken English language is greatly served.

Within the parameters defined by the free, protected and restricted syllables, stress is freely transferable between phonemes and syllables within connected speech.

5 This enables any word within an utterance to obtain prominence, or to stand out in some way within the phrase in the flow of speech. In a particular word in a phrase being made prominent or noticeable in sound and meaning, while the other words in the phrase do not gain as much prominence or do not stand out in the same way, the agreed upon meanings of all the words and the standardised or readily understood forms, functions, and structures in the phrase nonetheless remain constant.

10 For example, any word in the phrase "You are going!" may gain prominence or stand out in some fashion without negating the agreed upon meanings of any of the words in the phrase and without disturbing the standardised or readily understood forms, functions, and structures of the utterance. Moreover if the speaker were to make a particular word in the phrase especially prominent or noticeable then the cumulative meaning of the utterance would immediately change. This is because the word that gains prominence or stands out in the flow of speech can signify second order meaning.

15 Thus, for example, in addition to the agreed upon meanings of the words in the phrase being signified second order meanings may, in the first instance, also be signified according to the speaker's arrangement, composition and placement of stress and vocal prominence within the phrase.

For example:

20 "YOU are going!" makes the word "you" prominent and puts the emphasis on the subject, personally.

"You ARE going!" makes the word "are" prominent and would seem to affirm the subject's intention "to go", and assumes the possibility the topic has already been broached.

"You are GOing!" makes the word "going" prominent and focuses attention on the subject's act of "going".

25 It need be noted that it is not always the case that in order for a word to gain prominence or be noticeable within an utterance main stress must attend that word in order to signify second order meaning. For example, a word supported by only weak or secondary stress in a phrase may nonetheless gain prominence within a flow of connected speech that is dominated by heavy stress beats and high volume. The unstressed word could still gain prominence and stand out in the flow of speech and hence be capable of generating specific second order meaning.

30 Furthermore any word in a phrase may gain more particular qualities of prominence or be made more distinctly noticeable to the listener by way of the contours of variable tone and pitch and the various other sound and timing variations and contrasts that attend the phrase. These extra variable sound features of spoken phrases and sentences enable more complex and additional second order meanings to be signified to the listener. The signifiers of such complex and additional second order meanings would be built upon or around the variable stress inherent in the speaker's delivery of the words in the phrase had obtained and will most commonly introduce highly subjective ideas and meanings into the full quotient of meanings that the phrase can simultaneously support. Additional meanings of this nature would also be highly context dependent. Furthermore, variable sound imagery may signify various forms, functions or structures (such as, for example, changes in register, mood, the highlighting of case and the relationships between the subject, verb phrase and speaker).

40 For example, in particular contexts and depending on the sound and timing qualities of the speaker's variable sound imagery:

"YOU are going!" in placing the focus on the subject could also express the speaker's subjective opinion of the subject (e.g. enthusiasm or disgust);

45 "You ARE going!" in placing the attention on the subject's intention "to go" may also signify a curt imperative and not simply be a casual affirmation, and

"You are GOing!" in placing the attention on the subject's act of "going" may also express the speaker's personal feelings in regards to the subject's act of 'going' (e.g. regret or relief).

The flexibility of the phonetic and phonological systems of SE enables the variable sound imagery of individual phonemes, syllables and words to gain prominence in speech that signifies meaning. The permissive nature of the sound system of SE empowers the individual speaker with the vocal means by which he or she can choose which particular sound fragments, words, parts of speech, segments and passages of speech gain prominence within and between utterances and in what fashion prominence is to be achieved. This phonological freedom allows the individual speaker to productively signify second order meaning to the extent that the speaker's variable sound imagery obtained in the one spoken phrase may support several second order meanings all at once. Although in these cases the second order meanings emanate from the phrase's static first order meaning these can be extra meanings that are semantically in agreement with, independent of, or in opposition to the phrase's agreed upon 'first order' meaning. It need also be noted that second order meanings are meanings that realise no literal form as they are communicated by variable sound imagery and not words.

This aspect of SE greatly furthers the cause of the analytical language in its desire to express more complex meanings with reduced words, forms and structures.

#### - The Analytical Phonology and Word linking

The speaker is able to make any word, or words, within any utterance prominent through the use of variable sound imagery. The everyday habit of word linking aids the speaker in this cause for several important reasons, amongst them the following:

Variable sound imagery enables the speaker to differentiate one word from another in a linked, or partly linked, phrase. This is particularly necessary as words within utterances are very frequently linked to each other in natural speech. When word linking occurs the ends of individual words are changed or no longer distinct, nor do they necessarily need to be distinct, as spoken English is not a language system dependent upon inflections to communicate grammatical meaning. As well, the phonetic composition of the beginnings of words linked to preceding words is often effected. Moreover the phonetic content of linked English words may depart from their standard or readily understood phonetic definition – often radically -- at the boundaries between one linked word and the next. Most often word linking causes a reduction in the number of phonemes within utterances to occur, or in the replacement of the standard phonetic content of words with other phonemes (such as, for example, 'glides') which allow the linked words to be pronounced more easily and quickly by the speaker. The purpose of word linking can be regarded as a pragmatic means of reducing the forms and structures of utterances in that it preserves the time and energy of the speaker, thus serving the underlying analytic logic of spoken English;

In that word linking is a chief characteristic of natural English speech, the speaker is thus obliged to use variable sound imagery to make particular sounds, syllables and/or words distinct or comprehensible in particular ways that will communicate the agreed upon meaning of the word/s and enable the listener to quickly apprehend the speaker's meaning. The speaker depends upon variable sound imagery to highlight, distinguish, clarify and differentiate sounds, syllables, words and parts of speech from each other in linked and connected speech. Variable sound imagery may also help the speaker to organise and sequence the structures and functions of extended discourse. In this regard the listener depends upon the speaker using variable sound imagery in natural linked and connected speech, for without variable sound imagery the listener would be lost and unguided in a sea of undifferentiated verbiage;

In employing variable sound imagery in natural speech for the practical necessity of accommodating word linking the speaker is thus availed of the further opportunity of employing variable sound imagery for the communication of second order meanings. This is because the onus is placed on the speaker to decide which word or words within an utterance are to be made prominent and in what fashion and for what purpose, and the manner in which variable sound imagery may be used for the signification of various forms, functions and structures. This opportunity for signifying second order meaning is heightened by the fact that the majority of words in natural discourse are monosyllabic words which can be regarded as the 'free syllables' (that is, potentially capable of assuming any gradation and composition of 'stress' no matter how reduced or prominent).

- **The Analytical Phonology: The general principle of variability**

The immense effects of variable stress in SE, a 'variably timed' phonology, and the variable sound features of speech these engender, commission the speaker with an abundance of prosodic devices that may freely be used in speech for creating the variable sound imagery that signifies second order meaning. Furthermore, the speaker may work with other factors to generate variable sound imagery possessing meaning. For example, second order meaning could be obtained by speakers:

- varying the sound imagery of sounds and phonemes in the same words and phrases when repeated, or repeating back the same sound imagery using different words and phrases;
- varying the sound imagery of their own speech in order to create contrast, juxtaposition or counterpoint to meanings of the first or second order embedded within their own speech or with those of another speaker;
- using variable sound imagery to signify grammatical, linguistic and syntactical forms, functions, and structures without the need of verbal formulations, or with reduced verbal formulations;
- using sound and timing variation and contrast, however distinctive and subtle, to create 'sound metaphors' so as to express ideas, meanings and moods in conjunction with, independently of, or even oppositional to the lexical content of a particular utterance, such as, for example, to convey feelings and ideas of hesitancy, impatience, anxiety, or delight, etc;
- using variable sound imagery to communicate meanings of a sub-textual or abstract nature that may extend and expand across longer passages of speech and conversation;
- using variable sound imagery that contrasts, conflicts with, or colours, the facts of the immediate context in some fashion, or which plays tricks with the listener's expectations in regards to the kind of sound imagery that the listener expects to hear within a particular context or that customarily 'should' attend certain words and agreed upon meanings;
- combining any of these variations with other different legitimate variation/s and practices in order to create tension or contrast with the readily understood standard indicators of speech;
- inventing and negotiating new standard indicators of speech for certain idiolects, or simply for particular words and phrases, by which more variations that signify specific meanings may gain legitimacy within a particular speech community.

These examples, as listed here, should not be construed as limiting the number of ways a speaker may achieve prominence and variation in the sound imagery of connected speech for the purposes of signifying second order meaning. Rather, they illustrate the fact that a general principle of variability governs the phonology of SE because the number of different ways in which variable sound imagery signifying second order meanings could actually be obtained in acts of speech is almost without limit.

It can be said that in principle:

- the standard indicators of speech as understood by the particular community of SE speakers, plus

the variable sound imagery of these standard indicators of speech generated by the individual speaker and by other individual speakers within and/or without the immediate context, plus the variable and unpredictable compass of the immediate context of natural speech at its moment of utterance,

- 5 together, furnish SE discourse with the necessary linguistic constituents that facilitate the phonology's systems of internal tension, play and contrast by which variable sound imagery may obtain second order meaning in speech and upon which the entire system of the second order signification depends.

- **The Analytical Phonology and the Faculties of the Speaker and Listener in SE**

- 10 The listener can understand such meanings communicated by sound and timing variations because of the appreciable and audible contrasts they create within the stream of connected speech in ways that the fluent listener notices, registers, decodes and attempts to interpret for meaning. This is a subliminal process as the intention to signify, and the ability to interpret, meanings in the second order of signification need not be conscious.

- 15 For the decoding and interpreting of second order meaning the speaker and listener need standardised reference points by which to judge how much and in what ways sound imagery varies. For this they must rely on the standardised pronunciation of words, as the dictionaries or common usage define it, to gauge variations. Listeners and observers will also rely on commonly understood uses of pitch and tone in SE, in the interpretation of more complex second order meanings that the speaker's use of variable pitch, tone and stress may signify. The listener and observer also rely on their own knowledge of the speaker and the context of the conversation in which the variations occur to help him or her evaluate any second order meanings. It need be noted that because second order meanings can often be highly subjective so too is their possible interpretation by the listener.

- 25 Listeners and observers also rely on their own knowledge of English words and grammar and their own knowledge of the standardised indicators of the spoken English language, dialect, variety or idiolect to help them interpret second order meanings.

Amongst other things, the analytical phonology of SE requires both speaker and listener:

- having the highly specialised mental and physical faculties to formulate and generate second order meanings;
- 30 · possessing the complex of sophisticated and acute comparative and relative analytical skills and sensibilities needed to hear, recognise, register, measure and graduate sound and timing variations - however subtle, intricate and concurrent these may be - and with the ability to then interpret variable sound imagery for specific meaning;
- possessing an acute and abiding awareness of their immediate context which is axiomatic to the functioning of the second order as this dimension of the linguistic system is context dependent as the immediate context gives second order meaning motive. This 'context-awareness' must also be able to efficiently adapt to the ever changing contexts of daily discourse;
- 35 · having the various skills, faculties and sensibilities the SE language demands being properly co-ordinated in order to engage effectively in social communication.

- 40 For SE to realise its primary objective of increasing the communication of meanings with reduced and compressed forms and with no loss of efficiency- preferably, with greater efficiency – the relationship between interlocutors needs to be closer and more co-operative in SE than is the case between them in non-analytical languages. This is because the variable sound imagery of SE signifies meaning that is not agreed to meaning but rather is context-dependent meaning which is speaker generated and determined.

Accordingly interlocutors need to possess, between them, the ways and means – physical and mental - to readily establish systems and modes of communication with each other that will enable the free flow of second order meanings to occur in the immediate context of spoken discourse and in ways that best serve efficiency and conserve the time and energy of those involved.

5 Such faculties are among the fundamentals that the native English language speaker must first acquire in order for this to occur. They also typify the kind of skills and faculties that speakers of languages other than English must learn to acquire in tandem with those they already possess, and which voice recognition technologies must simulate, if effective communication in SE is to eventuate.

#### 10 - **Spoken English: a new definition & conceptual framework**

It is believed by the inventor that both Saussure's original construct of the language sign and modern Applied Linguistics long held definition of language are incomplete definitions of the spoken English language and its analytical phonology.

15 Saussure's concept of the language sign and the orthodox view of human language describe the first order of signification only.

It can be said of a language system with only one order of signification that:

standardised sounds and phonemes = words which are arbitrary symbols possessing no more than their agreed-upon meanings.

20

Therefore words, as so defined, are the singular currency of linguistic signification available to the speaker for the production of meaning. In some first order only languages further meaning may be obtained through the use of the suprasegmental and prosodic features of speech where these are possible, available and permissible but strictly under the proviso that their use does not interfere with the primary purpose of the one fixed and central timing principle that controls connected speech: to organise and control the flow of connected speech in a way that ensures the agreed-upon meanings of the words are protected and remain fixed and unvaried.

25

Spoken English, as opposed to the conventional language systems defined by Saussure, possesses active and highly productive first and second orders of signification. Because the language has two orders, and not one, the relationship between the first and second orders of signification alters the definition of the first order.

30

In the analytical phonology it can be said that within its first order of signification:

standardised sounds and phonemes = words which are arbitrary symbols that possess agreed-upon meanings.

35

Within the language's second order of signification it can be said that:

in the speakers' use of sound and timing variations and contrasts, variable sound images of standard words are obtained in speech in a way that does not negate the first order signification but, moreover, systematically generates the signifiers of further or new meaning, or enables the speaker to express meaning using reduced forms, functions and structures, ie the second order of meaning.

40

Within the second order of signification the basic signifiers in the production of meaning in speech are phonemes which construct spoken phrases which are the basic units of second order meaning. Hence the SE phrase in everyday connected speech is understood by the native speaker as being both a lexical-grammatical entity while simultaneously being a – potentially highly – variable sound entity.

45

Therefore, it can be said that words within the analytical phonology of SE possess two values in meaning:

- the first value as an arbitrary symbol possessing a static agreed to meaning but with the potential to expand to concurrently gain
- 5 a second variable value in meaning that the word may obtain in speech as a subjective and relative symbol. 'Subjective' in the sense that it is meaning that is speaker generated and context dependent, with the speaker and the facts of the immediate context giving second order meaning its motive and sense. 'Relative' in the sense that the second variable value in meaning is 'relative' to the word's first order meaning and its standard indicators of speech. It is also 'relative' to the variable sound imagery
- 10 the word obtains in speech and, in certain instances, 'relative' to the variable sound imagery the same word or other words may also obtain in speech.

It need be noted that the word's first value, its static agreed upon meaning, is capable of expanding to simultaneously support more than one variable value in meaning, this being executed in the word's moment of utterance. For example, in a certain context, a speaker may employ variable sound imagery in response to a

15 persistent 'yes/no' question. The speaker may reply: 'Yes' to the question, but by way of variable sound imagery clearly signify 'No', while at the same time be signifying the additional idea: 'And don't keep asking me this question all the time'. Here the listener is presented with a one-word answer that supports three meanings simultaneously: one agreed upon meaning plus two second order meanings. It need also be noted that a word's second variable value in meaning may be in semantic agreement with, complementary to,

20 independent of, or in opposition to, its agreed upon value in meaning and/or different second order meanings the same word may obtain in speech.

The fact that all English words may enjoy two potential values in meanings in speech lends SE an inestimable vocabulary for the purposes of 'making meaning'. Its corpus can be regarded as being two dimensional. One dimension is comprised of the 'hard' words of the first order of signification where words are

25 arbitrary symbols that possess static agreed upon meanings. Its second dimension is 'virtual' being comprised of the potential second order meanings that the arbitrary symbols may realise by way of their variable sound imagery, with these arbitrary symbols awaiting a speaker and a 'live' context for their formulation or reformulation of meaning within the second order of signification.

Among the hallmarks of the modern spoken English language which characterise its underlying logic and

30 unity of purpose and that would distinguish it from the syllable and stress timed language systems are the following:

- the general principle of variability governing the phonology and sound imagery of spoken discourse which generates the variable sound imagery of natural speech. This invests all words in connected speech with two potential values in meaning: one standardised, one variable,
- 35 lending the language an inexhaustible vocabulary of potential meanings;
- reductionism: variable sound imagery may signify various grammatical, syntactical and linguistic forms, functions, and structures in natural speech;
- compactness achieved through the compression of meaning in the second order of signification, and other forms of reductionism obtainable in the first order (such as, for example, the use of
- 40 contractions, ellipsis, slang etc);
- diversity: exemplified by the language's copious 'multicultural' corpus, and the many varieties and idiolects of spoken English that the language engenders and invents;
- an expansionist drive: seen in the language's capacity to acquire new loan words from other languages, a ravenous, on-going process that its phonological system readily accepts when
- 45 exotic words come to be placed into connected speech. Evincing, also, by the ever expanding

parameters of legitimate sound variability tolerated in modern and popular varieties over which no official high arbiter of 'correct speech' or language planning presides but which mass media dominates;

- 5       - the ability to self-reform and adapt, in that the progressive analytical objective of the language will generally prevail over tradition and convention - in nearly all varieties of SE bar genres such as Received Pronunciation - when the two are in conflict. RP and other recognised stable genres of SE usually change only gradually - or, in the case of genres such as 'Network-American-English' these can sometimes change more rapidly via mass communication - all providing the necessary standardised indicators and conventional modes of speech which
- 10       furnish discourse with the stable linguistic constituents that make the internal tension, play and contrast, which engenders variable sound imagery, intelligible;
- an ethos of individualism as the system of second order signification is highly speaker-centred;
- an ethos of indifferent equality between its speakers, exemplified by the practical necessity for the uniquely close and co-operative relationship between interlocutors that exists, irrespective of
- 15       differences in their age and status. In general, SE comfortably accommodates and proliferates informal and popular genres of speech;
- an uncommon 'poetics' of discourse, exemplified by the wealth of subjective, psychological and sub-textual meanings that may resonate through the variable sound imagery of ordinary words and phrases signifying second order meanings which realise no literal form;
- 20       - a high degree of linguistic involvement in that both systems of signification are mutually entwined and dependent, it being highly unnatural to conduct 'first order only' SE conversation, lest the primary analytical objective be lost. Even so, operating with only one order is inefficient and problematical as an absence of appreciably variable sound imagery in everyday SE disorients and disengages the listener; and
- 25       - efficiency, SE being the natural mother tongue of capitalism.

In short, SE can be regarded as a liberal-democratic institution in the classical sense of the term.

Turning now to the preferred embodiments illustrated in FIGS 3 and 4, there is shown in FIG 3 a flow chart of a method of recognising speech in accordance with the invention.

- 30       The method of recognising natural English speech (which consists of words having syllables and phonemes) includes assigning to SE words and speech a first order of signification (12) which includes words having standardised indicators possessing agreed meanings, and speech possessing standardised or readily understood grammatical, syntactical and linguistic forms, functions, and structures independent of the speaker. It also includes designating the syllables in words as being protected, restricted or free syllables which assigns a
- 35       potential variability to each syllable that may be obtained in connected speech. A second order of signification is also assigned to SE words and speech (13), this second order having words possessing variable indicators which have meanings, and speech possessing grammatical, syntactical and linguistic forms, functions, and structures which are generated by the speaker's use of variable sound imagery and that are dependent on the context of the word/s and utterance/s in the flow of connected speech.

- 40       These variable indicators include the pronunciation (14) of phonemes, syllables and words in the speech with the syllables categorised as being either free syllables (15), protected syllables (16) or restricted syllables (17) according to the syllables' variable indicators which include the key phonological functions of speech (18) such as pitch, volume, tone, duration, rhythm and tempo as well as including other suprasegmental features of speech (19) such as, speed of delivery, enunciation, pausing, phrasing and word linking.



The words are then analysed (20) in accordance with the first and second orders of signification. Integral to the process of analysis is the defining of the pertinent facts of the immediate context of the words in speech. Factors such as: **When? Where? Why?** is the conversation occurring? **How** is the conversation occurring (mode of exchange: face-to-face, by phone, via technology etc)? **What** is the social purpose or the business of the exchange and what events have led up to the conversation that have relevance to the immediate context? **Who** are the participants in the exchange, what is their relationship, and what are their manner, mood and temper of mind? This information may be either operator-dependent, or may be generated or anticipated by the technology within pre-defined parameters and contexts.

In use the method is implemented by recording speech spoken by a speaker (11), analysing the recorded speech as above, and then indicating to the speaker (21) the meanings of the variable indicators of the recorded speech. The speaker then designates or affirms (22) the meanings of the variable indicators which have been indicated to him or her. Data representative of analysed words for which the meanings of the variable indicators have been designated or affirmed is then stored (23) in storage means for subsequent transformation to another format (24), such as another language, WP text etc.

FIG 4 is a schematic block diagram illustrating a system for recognising speech such as SE consisting of words having syllables and phonemes and speech possessing grammatical, syntactical and linguistic forms, functions, and structures. The system has a recorder 31 for recording speech spoken by a speaker 35. Assigning means 32 assigns a first order of signification to a word, the first order of signification includes words having standardised indicators having agreed meanings, and speech possessing standardised or readily understood grammatical, syntactical and linguistic forms, functions, and structures independent of the speaker. Assigning means 33 then assigns the facts of the immediate context, as described above, and a second order meaning to the word/s and speech. The constituent elements of the second order of signification are the same as those described above.

Indicating means 34 indicates to speaker 35 the meanings of the variable indicators in the speech which has been recorded. Designation means 36 are operable by speaker 35 to designate or affirm the meanings of the variable indicators which have been indicated by indicating means 34 for subsequent transformation by transforming means 39 to output means 40 such as, for example, a printer of WP text, a computer controlled human voice simulator etc.

Analysing means 37 analyses words or speech in accordance with the first and second orders of signification and data representative of the analysed words and speech, for which the meanings of the variable indicators have been designated or affirmed, is stored in storage means 38.

Assignment means 32 and 33, analysing means 37, storage means 38 and transforming means 39 are embodied in a suitably programmed computer 41, the peripherals to which include recorder 31 such as a microphone, indicating means 34 (screen), designating means 36 (keyboard or mouse) and output means 40 (printer, screen, speaker etc).

It will be obvious to those skilled in the art that there are numerous ways in which the present invention can be practised in VRT, that suitably skilled programmers can write software embodying the teachings of this invention in its various applications to VRT, and that suitably configured and programmed computing systems can be utilised to practice the invention.

Whilst the invention in its application to VRT can be adequately performed by those skilled in the art on the basis of the description thus far provided, a number of features and aspects relating to implementation of the invention are further provided. These are not to be construed as limiting on the scope of the invention.

With reference to the preferred VRT implementation of the present invention:-

The system is designed for application at the level of the individual speaker or the individual operator who will use the system.

The system first defines the standardised value of words according to their phonetic content and agreed-upon meaning as is defined by a particular community of native English language speakers. Also, the system first defines the standardised or readily understood grammatical, syntactical and linguistic forms, functions, and structures representative of the variety of SE understood by the operator.

5 The individual speaker has the opportunity for direct input into defining or editing what the standardised "sound" values and meanings of arbitrary symbols, and the standardised forms, functions, and structures of his or her speech, are to be, although specific software packages targeted for identified English language "speech communities" should already be cognisant of the various standardised indicators prevalent within that particular group. This process initially establishes the essential reference point  
10 which variable sound imagery can be measured against and evaluated when words and speech are used in natural discourse.

Once the standardised values of English words and speech have been defined by the system with the speaker's supervision and/or input, the system then defines the second order meanings of words and speech.

15 The system samples the speech of the individual speaker and operator at its most natural and spontaneous, and in long durations. This is preferably done in software generated contexts such as games, conversations and particular scenarios with which the individual speaker and operator is familiar and orientated. This allows the individual to produce second order meanings within a defined or known context in an unrehearsed and spontaneous sampling of his or her everyday connected speech, allowing  
20 the individual to demonstrate a full and natural range of his or her vocal and pronunciation styles, emotions and registers. The individual's particular and authentic qualities and manner of speech may thus be recorded.

As the system samples and records the speaker's manner and qualities of speech, it first measures and qualifies his or her specific parameters of sound and timing variability, paying particular  
25 attention to the specific factors outlined in the preferred embodiment above.

Upon establishing the general parameters of variability evinced in the operator's natural and spontaneous speech, the system notes, measures and defines any variations and contrasts that arise, using the reference point of the standardised sound values.

30 The system notes the individual's particular tendencies, habits and patternings of pronunciation and voice in natural spontaneous speech and then alerts the operator to sound variations and contrasts evident in the operator's samplings of natural connected speech (as compared to the standardised values of words and speech already established).

35 Within the defined facts of the immediate context the system then allows the speaker to define the particular second order meanings that such variations and contrasts may signify. For instance, do certain combinations of sound and timing variations (say the elongating or reducing of the vowel sound in the same word when repeated) mean something? Or do they signify second order meanings which the speaker had not consciously intended to generate but nonetheless now need to be consciously considered by him or her in light of the facts of the immediate context? Do clusters of contrasts noted by the sampling process in the operator's natural speech and which cause individual phonemes, syllables  
40 and words to noticeably vary and deviate from the standardised norm, possess second order meaning, and if so, what? An emotion, irony, a subtext of some kind? Do they signify a particular form, function, or structure? Often an individual speaker's peculiar habits and manner of pronunciation signify a personal meaning, mood or theme, irrespective of what the actual words in the phrase might be, but are still relative to and dependent on the immediate context. In these instances the system offers the operator

the choice of identifying and labelling such recurring variations and contrasts for any specific meanings they may possess within the frame of reference of the immediate context.

The definition of second order meanings is preferably effected by systems of menus, sub-menus and options over which the trained operator has executive control. Individual operators are allowed to define what sound and timing variations and contrasts in their speech signify or mean.

In order to correctly define second order meanings, the system preferably understands the grammatical systems and principles of the spoken word: for example, such things as word order, syntactical formulas, strategies of "reductionism" in the first order (such as ellipsis and contractions), grammatical categories etc, so as to divine the grammatical rationale behind the speaker's use of variation and contrast.

In summary, in effectively communicating with the native English speaker, the system multi-tasks in the sense of simultaneously being aware of the first order of signification and the processes of the second order of signification. It is here that meanings are routinely signified by variations and contrasts in the individual speaker's speech and the system requires reference back to the first order of signification, a knowledge of the context, and a knowledge of the individual speaker, in order that second order meanings can be properly and fully understood. In other words, the system does what individual English language speakers constantly do in order to understand and appreciate the second order meanings of natural, speech which the variable sound imagery of everyday spoken English generates.

It will be appreciated that the system in accordance with the present invention can effect a number of desirable and advantageous outcomes including:-

Translating both orders of signification that exist in the spoken English language into word, and vice-versa, without negating or dulling the layers of second order meanings that the prosodic features of the spoken word naturally support and signify in speech.

Allowing individual English language speakers and operators to more effectively communicate with other English language speakers and operators.

Allowing individual English language speakers to communicate with speakers of other languages in both speech and by the written word, such that the second order signification present only in speech will not be lost or nullified when translated or put into the written word.

Allowing individual English language speakers who are hearing impaired or deaf to appreciate second order signification and second order meanings that are encoded within natural speech through the written word.

Allowing individual English language speakers who are sight impaired or blind to appreciate second order signification and second order meanings encoded within the written word and which can be translated back into natural speech.

The present invention can also be used to teach how to speak a language. In use this method includes:- assigning first and second orders of signification to words and speech, the words having syllables and phonemes, the first order of signification including standardised indicators having agreed meanings, and speech possessing standardised or readily understood grammatical, syntactical and linguistic forms, functions, and structures independent of the speaker and the second order of signification including variable indicators having meanings, and speech possessing grammatical, syntactical and linguistic forms, functions, and structures which are generated by the speaker's use of variable sound imagery and that are dependent on the context of the word in the flow of connected speech, and

practicing speaking using different variable indicators in the second order of signification.

When teaching how to speak the English language in schools in English speaking countries information concerning the analytical nature of the language should be included in mainstream English curriculums.

A number of aspects associated with the present invention are emphasised in the curriculums. These include:-

Providing and teaching a proper and appropriate explanation of the nature, practices and principles of the analytical phonology of spoken English.

5 Explaining to children, from the moment they learn to read and write, the nature of the separation of the language's written and spoken systems at the level of phonemes and their written alphabetical symbols. (Then relying on rote and the remarkable powers of retention and memory children possess in order to teach spelling - as there is no readily understandable connection between the two systems at the level of phonemes and their written alphabetical symbols which can be taught or learnt).

10 Encouraging, nurturing and fostering the expressive powers of SE's second order of signification in the individual child in formal genres of speech, as well as in informal or popular genres. Standardisation theories and strategies of speech should not be imposed on children, as this runs against the analytical purpose of SE which encourages individualism and inventiveness in speech to communicate as much meaning as possible in the shortest forms and structures obtainable. Children are aware of the expressive powers of second order signification. The various popular genres of speech children adopt within their peer groups, which for the child is the most dominant linguistic influence, are rich in second order meanings.

15 The teaching of ESL is similar to that above and much of what is taught to native English language speakers regarding the analytical nature of modern English is also taught to those learning English as a non-mother tongue. One way this can be achieved is by conveying the information to the ESL learner in his or her own language to provide a grounding in the basic principles of the system before learning commences.

20 The many advantages of the present invention over traditional speech recognition systems and methods and the various applications thereof will already be apparent to the addressee. In brief summary these advantages stem from the central fact that SE is an analytical phonology with first and second orders of meaning or signification.

25 Thus SE is a "speaker-centred" language with enormous executive power delegated to the individual speaker in producing second order meanings according to the individualistic way they may vary the sound and timing qualities of words when used in everyday connected speech.

30 Present VRT package programs require the operator to painstakingly sample speech word by word. This sees English words as purely arbitrary symbols with no regard to the words' potential second relative or subjective value that is clearly obtainable in natural connected speech. Existing VRT systems record the word's static agreed-upon meaning in the first order of signification where words as arbitrary symbols possess no more than this and generally have a standardised pronunciation. Consequently for the computer in existing VRT packages to familiarise itself with the particular operator and recognise the operator's manner of speech, the onus falls heavily on the operator not to vary his or her pronunciation of the sampled words in any substantial way when the system is later in use, lest the machine be unable to recognise the words correctly.

35 The failure of known VRT systems to come to grips with the second order of signification is believed by the inventor to have prevented computer science from developing the technologies to enable communication with native English speakers at a level above the somewhat robotic-like current speech level.

40 The invention is also applicable in telecommunications where existing recorded computer generated voices repeat back numbers to the telephone customer. These can be modified to sound more like natural connected speech.

It will of course be realised that whilst the above has been given by way of an illustrative example of this invention, all such and other modifications and variations hereto, as would be apparent to persons skilled in the art, are deemed to fall within the broad scope and ambit of this invention as is herein set forth.

**The Claims defining the Invention are as follows:-**

1. A method of recognising speech consisting of words having syllables and phonemes, said method including:-
  - assigning first and second orders of signification to a word;
  - wherein said first order of signification includes standardised indicators having agreed meanings independent of the speaker, and said second order of signification includes variable indicators having meanings which are generated by the speaker and are dependent on the context of the word in the flow of connected speech.
2. A method of recognising speech as claimed in claim 1, wherein the speech is spoken English.
3. A method of recognising speech as claimed in claim 2 and including:-
  - assigning first and second orders of signification to language;
  - wherein said first order of signification assigned to language includes forms, functions and structures independent of the speaker, and said second order of signification assigned to language includes variable forms, functions and structures which are generated by the speaker and are dependent on the context of the word(s) and/or utterance(s) in the flow of connected speech.
4. A method of recognising speech as claimed in claim 3, and including:-
  - analysing said word(s) and/or language in accordance with said first and second orders of signification.
5. A method of recognising speech as claimed in claim 4, wherein said variable indicators include the pronunciation of phonemes, syllables and words in the speech.
6. A method of recognising speech as claimed in claim 4, wherein said variable indicators include features of speech such as variations in pitch, tone, harmonic content, volume, duration, rhythm, tempo and the rate of syllables spoken per unit time.
7. A method of recognising speech as claimed in claim 4, wherein said variable indicators include other suprasegmental or prosodic features of speech such as variations in the speed of delivery, variations in enunciation, variations in pausing, variations in phrasing and variations in word linking.
8. A method of recognising speech as claimed in claim 4, wherein said variable indicators include variable forms, functions and structures which are communicated by the speaker by way of variable sound imagery in speech.
9. A method of recognising speech as claimed in claim 4, wherein said variable indications include the facts of the immediate context pertaining to the words in the flow of connected speech.
10. A method of recognising speech as claimed in claim 5, wherein said syllables are categorised as being either free syllables, protected syllables or restricted syllables.
11. A method of recognising speech as claimed in claim 3 and including:-
  - recording speech spoken by a speaker;

indicating to the speaker the meanings of the variable indicators of the recorded speech, and designating or affirming the meanings of the variable indicators indicated to the speaker.

12. A method of recognising speech as claimed in claim 8 and including:-  
storing data representative of analysed words for which the meanings of the variable indicators have been designated or affirmed.
13. A method of recognising speech consisting of words having syllables and phonemes, the method including:-  
assigning first and second orders of signification to words and language;  
wherein the first order of signification includes words and language having standardised indicators having agreed meanings and forms, functions and structures independent of the speaker, and the second order of signification includes words and language possessing variable indicators signifying variable meanings and forms, functions and structures which are generated by the speaker and which are dependent on the context of the words and utterance/s in the flow of connected speech.
14. A system for recognising speech consisting of words having syllables and phonemes, said system including:-  
recording means for recording speech spoken by a speaker;  
means for assigning a first order of signification to a word, said first order of signification including standardised indicators having agreed meanings independent of the speaker;  
means for assigning a second order of signification to a word, said second order of signification including variable indicators having meanings which are generated by the speaker and are dependent on the context of the word in the flow of connected speech;  
indicating means for indicating to a speaker the meanings of the variable indicators of the recorded speech, and  
designation means whereby a speaker designates or affirms the meanings of the variable indicators indicated by the indicating means.
15. A system for recognising speech consisting of words having syllables and phonemes, the system including:-  
recording means for recording speech spoken by a speaker;  
means for assigning a first order of signification to words and speech, the first order of signification including standardised indicators having agreed meanings, and standardised forms, functions and structures independent of the speaker;  
means for assigning a second order of signification to words and language, the second order of signification including variable indicators possessing meanings, and forms, functions and structures which are generated by the speaker and are dependent on the context of the word in the flow of connected speech;  
indicating means for indicating to a speaker the meanings, and forms, functions and structures of the variable indicators of the recorded speech, and  
designation means whereby a speaker designates or affirms the meanings, and forms, functions and structures of the variable indicators indicated by the indicating means.
16. A system for recognising speech as claimed in claim 15, and including:-

analysing means for analysing said words and speech in accordance with said first and second orders of signification.

17. A system for recognising speech as claimed in claim 15, and including:-

storage means for storing data representative of analysed words for which the meanings and forms, functions and structures of the variable indicators have been designated or affirmed.

18. A method of teaching how to speak a language, said method including:-

assigning first and second orders of signification to a word, the word having syllables and phonemes, said first order of signification including standardised indicators having agreed meanings independent of the speaker and said second order of signification including variable indicators having meanings which are generated by the speaker and are dependent on the context of the word(s) in the flow of connected speech, and practicing speaking using different variable indicators in said second order of signification.

19. A method of teaching how to speak a language as claimed in claim 18, and including:-

assigning first and second orders of signification to language, said first order of signification including forms, functions and structures independent of the speaker and said second order of signification including forms, functions and structures which are generated by the speaker and are dependent on the context of the word(s) in the flow of connected speech.

20. A method of teaching how to speak a language as claimed in claim 19, said method including:-

practising speaking in order to acquire the preferred respiratory, cognitive and vocal skills, and the preferred skills of physical and mental co-ordination, in using variable indicators in the second order of signification.

21. A method of teaching how to speak a language as claimed in claim 20, said method including:-

analysing connected speech that exhibits different variable indicators in the second order of signification for the purposes of recognising and evaluating the speech for meaning.


22. A method of teaching how to speak a language as claimed in claim 19, wherein said language is spoken English.

23. A method of teaching how to speak a language as claimed in 22, wherein English is taught as a first language.

24. A method of teaching how to speak a language as claimed in 22, wherein English is taught as a second language.



1/5



disapPOINTing

FIG 1A

standardised  
word stressing

*disappointing*

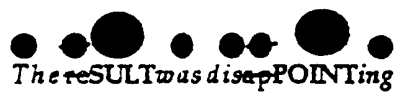
*dis, ing* = *free* syllables

*ap* = *restricted* syllable

*point* = *protected* syllable

FIG 1B

potentially variable  
word stressing



The reSULT was disapPOINTing

FIG 1C



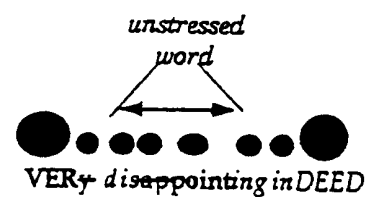
A DISappointing reSULT

FIG 1D



THAT'S DISappointING!

FIG 1E



unstressed  
word

VERY disappointing in DEED

FIG 1F

2/5


Rising A yes or no question: *Are you hungry? Now?*  
 A clarifying question: *What happened then?*  
 Please repeat: *Sorry? Pardon? Again?*  
 Involved feedback: *Really?*

FIG 2A


Low rising Reading items from a list: more to come  
 *Sydney, Canberra...*  
 Neutral feedback: *Uh-huh, Mmmm*

FIG 2B


Level Monotone.  
 Bored or indifferent: *Hello.*  
 Ironic, sarcastic: *Sorry. Really?*  
 Routine: *Next please.*

FIG 2C


Rising-falling Exclamatory: *Fabulous! Awesome!*  
 Opinion: *Well I think...*  
 Dramatic: *Now!*  
 Conversation markers: *Anyway...*

FIG 2D**KEY**

/free syllables  
 restricted syllables  
 protected syllables

3/5

Falling-rising

Conjunctions: *However, meanwhile,*  
*nonetheless* to show contrasting ideas  
 Some modal verbs: *Could be, should be,*  
*can't be, might be, wouldn't be* etc

FIG 2EFalling

Last item on a list to show completion  
 of ideas: *Sydney, Canberra, Melbourne*

A *wh* question: *Where do you live?*

FIG 2F

Declaratives: *I live in Sydney.*

Apology: *Sorry, Pardon*

Sharp falling

Strong Apology: *Sorry, Pardon*

Imperative: *Now!*

Declarative: *Goodbye.*

FIG 2G

*Did you murder your wife?*

FIG 2H

Speaker A: *I'm going to stop drinking.*

FIG 2I

Speaker B: *I'm sure you will.*

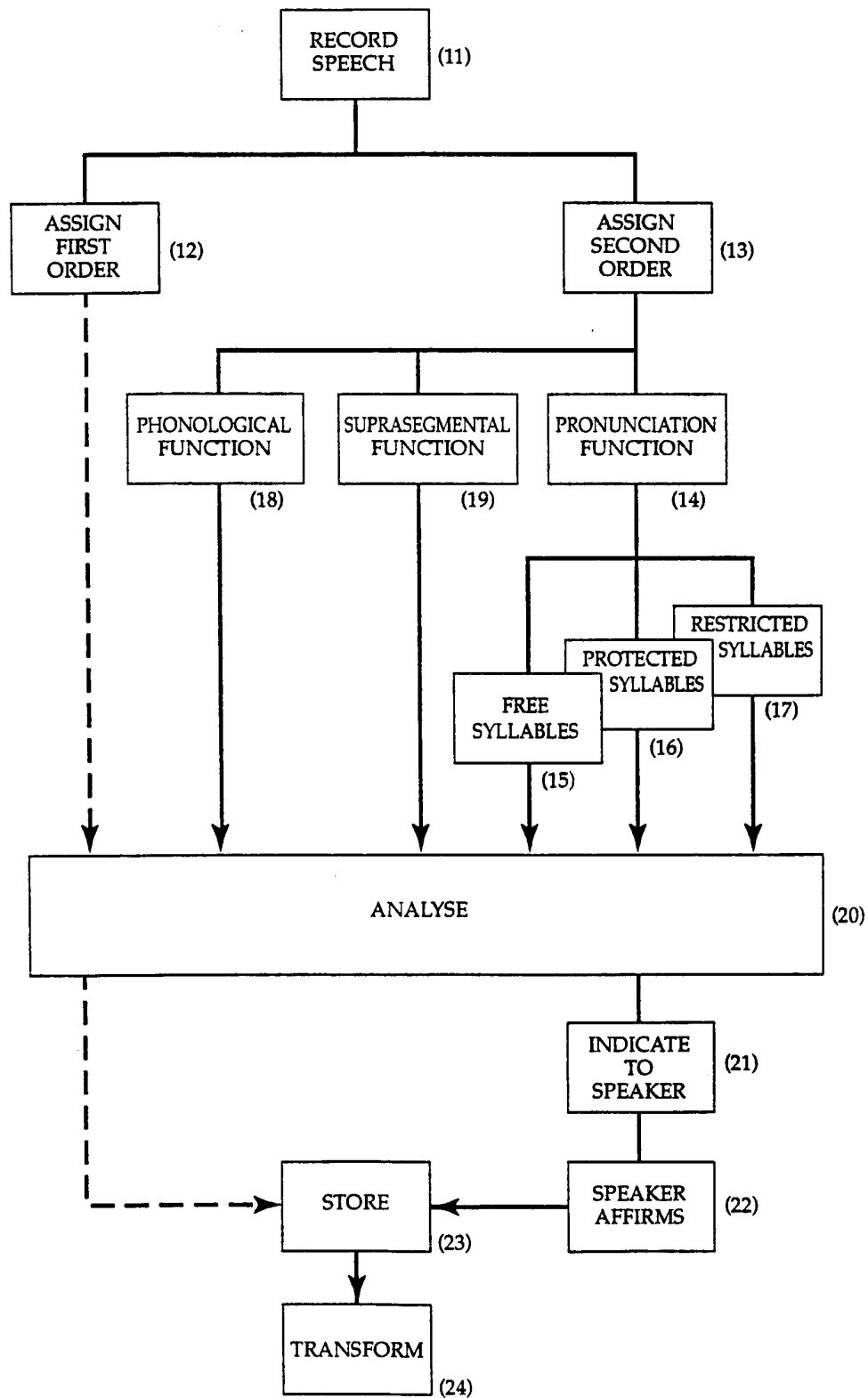
FIG 2J**KEY**

*free syllables*

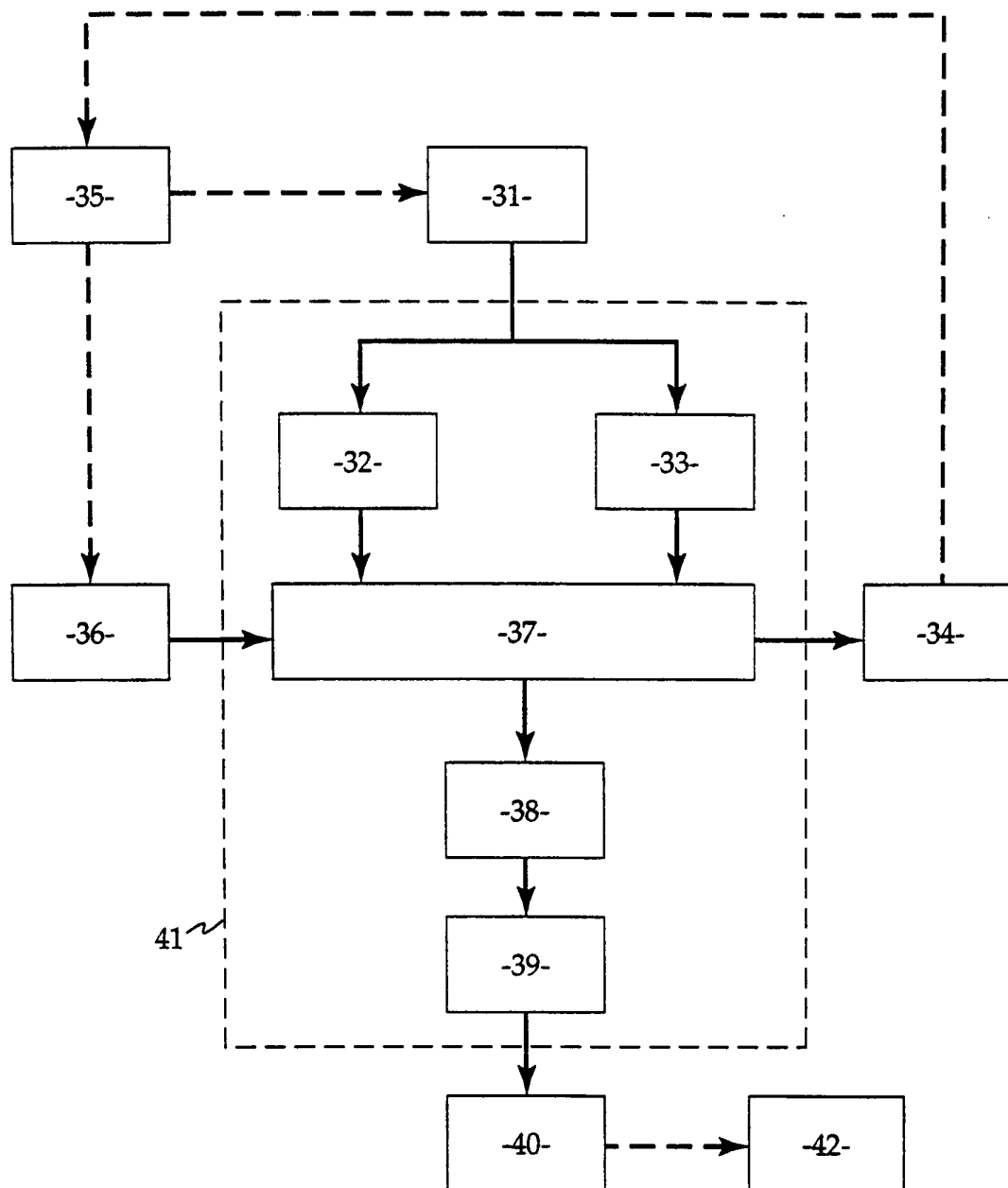
*restricted syllables*

*protected syllables*

4/5


*Fig. 3.*

5/5

*Fig. 4.*

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/AU00/00817

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>												
Int. Cl. <sup>7</sup> : G10L 15/00, 11/00; G09B 19/04, 19/06												
According to International Patent Classification (IPC) or to both national classification and IPC												
<b>B. FIELDS SEARCHED</b>												
Minimum documentation searched (classification system followed by classification symbols) IPC: G10L, G09B 19/04, 19/06												
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 practicable, search terms used) WPAT: G10L/IC AND (stress OR context OR pitch OR intonation) AND "speech recognition"												
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>												
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.										
X	US-5806033-A (LYBERG) 8 September 1998 & US-5752227-A (LYBERG) 12 May 1998 & US-5677992-A (LYBERG) 14 October 1997 & US-5546500-A (LYBERG) 13 August 1996 See all four documents.	1 to 7, 11, 13 to 17										
X	US-5636325-A (FARRETT) 3 June 1997 See column 6 line 32 to 65, column 7 line 48 to column 8 line 67	1 to 7, 13										
X	US-4802223-A (LIN et al.) 31 January 1989 See column 2 lines 16 to 32, and column 4 line 32 to column 7 line 29.	1 to 4, 6										
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C <input checked="" type="checkbox"/> See patent family annex												
<p>* Special categories of cited documents:</p> <table border="0"> <tr> <td>"A" document defining the general state of the art which is not considered to be of particular relevance</td> <td>"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</td> </tr> <tr> <td>"E" earlier application or patent but published on or after the international filing date</td> <td>"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</td> </tr> <tr> <td>"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)</td> <td>"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</td> </tr> <tr> <td>"O" document referring to an oral disclosure, use, exhibition or other means</td> <td>"&amp;" document member of the same patent family</td> </tr> <tr> <td>"P" document published prior to the international filing date but later than the priority date claimed</td> <td></td> </tr> </table>			"A" document defining the general state of the art which is not considered to be of particular relevance	"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	"E" earlier application or patent but published on or after the international filing date	"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	"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)	"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	"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family	"P" document published prior to the international filing date but later than the priority date claimed	
"A" document defining the general state of the art which is not considered to be of particular relevance	"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											
"E" earlier application or patent but published on or after the international filing date	"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											
"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)	"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											
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family											
"P" document published prior to the international filing date but later than the priority date claimed												
Date of the actual completion of the international search 19 September 2000		Date of mailing of the international search report 25 SEP 2000										
Name and mailing address of the ISA/AU AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustalia.gov.au Facsimile No. (02) 6285 3929		Authorized officer  MICHAEL HARDY Telephone No: (02) 6283 2547										

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/AU00/00817

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
E, A	US-6109923-A (ROTHENBERG) 29 August 2000 See abstract	18 to 24
A	US-5475796-A (IWATA) 12 December 1995 See abstract and figure 1.	

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

International application No.  
**PCT/AU00/00817**

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report				Patent Family Member			
US	5677992	EP	664537	JP	7181997	SE	9303623
US	5546500	DE	69420614	EP	624865	JP	6332494
		SE	9301596				
US	5806033	EP	749109	JP	9022297	NO	962463
		SE	9502202				
US	5752227	EP	683483	JP	8050498	SE	9401613
US	5636325	NIL					
US	4802223	NIL					
US	5475796	JP	5333892				
US	6109923	NIL					
END OF ANNEX							