



(72) KIRKSEY, WILLIAM E., US

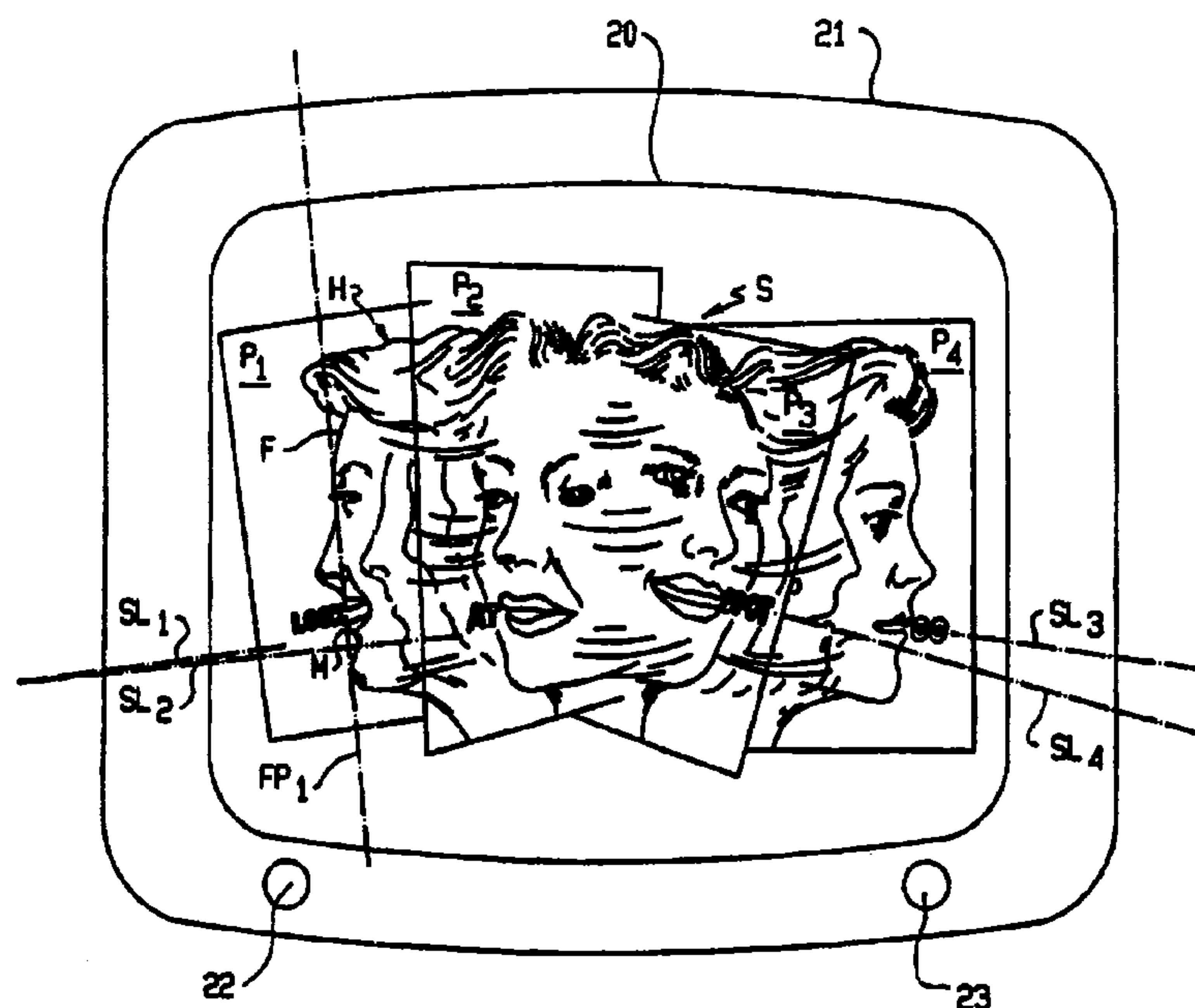
(71) READSPEAK, INC., US

(51) Int.Cl.⁶ G09B 5/00

(30) 1997/04/25 (08/845,863) US

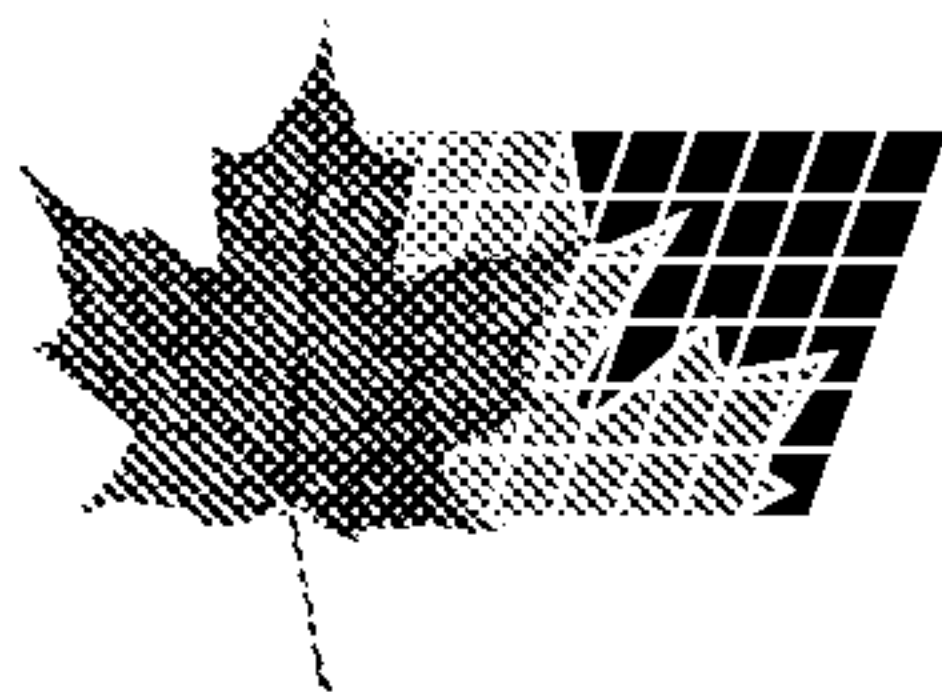
(54) **PROCEDE ET SYSTEME DE PRODUCTION D'UNE PIECE AUDIOVISUELLE COMPORTANT UNE SERIE DE SYMBOLES VISUELS DE MOTS ASSOCIES A DES ENONCIATIONS ORALES ET PIECE AUDIOVISUELLE CORRESPONDANTE**

(54) **METHOD AND SYSTEM FOR MAKING AN AUDIO-VISUAL WORK WITH A SERIES OF VISUAL WORD SYMBOLS COORDINATED WITH ORAL WORD UTTERANCES AND SUCH AUDIO-VISUAL WORK**



(57) L'invention concerne une pièce audiovisuelle et son procédé de création, laquelle pièce comporte, sur ses images, des mentions écrites, de manière qu'au fur et à mesure que l'utilisateur entend chacun des mots ou autre énonciation, une mention écrite à associer au son entendu apparaît alors, en évidence ou non, de sorte que, dans l'esprit du spectateur-auditeur original, la présentation future de l'énoncé évoquera les mentions

(57) An audio-visual work and method of its creation which work has writings placed on the pictures of the work so that as each word or other utterance is heard a writing to be associated with the hearing is coordinated with seeing by emphasis or otherwise of the writing such that the future presentation of either the utterance or writing shall evoke the other in the mind of the original viewer-listener. Each word will when appropriate appear



(21) (A1) **2,287,315**
(86) 1998/04/25
(87) 1998/11/05

écrites, et vice versa. Chaque mot apparaîtra au moment approprié sous une présentation de lecture adjacente à la bouche (M) du locuteur (S). L'affichage de la pièce peut se faire de façon linéaire ou sur commande informatique du spectateur-auditeur, et peut être accompagné d'autres supports éducatifs. Il est possible de procéder à une utilisation sélective de groupes de mots mis en évidence ou non.

in a legible perspective adjacent to the mouth (M) of the utterer (S). The work can be displayed linearly or under computer control of the viewer-listener along with additional educational materials. Selected use of groups of words highlighted or not may be used.



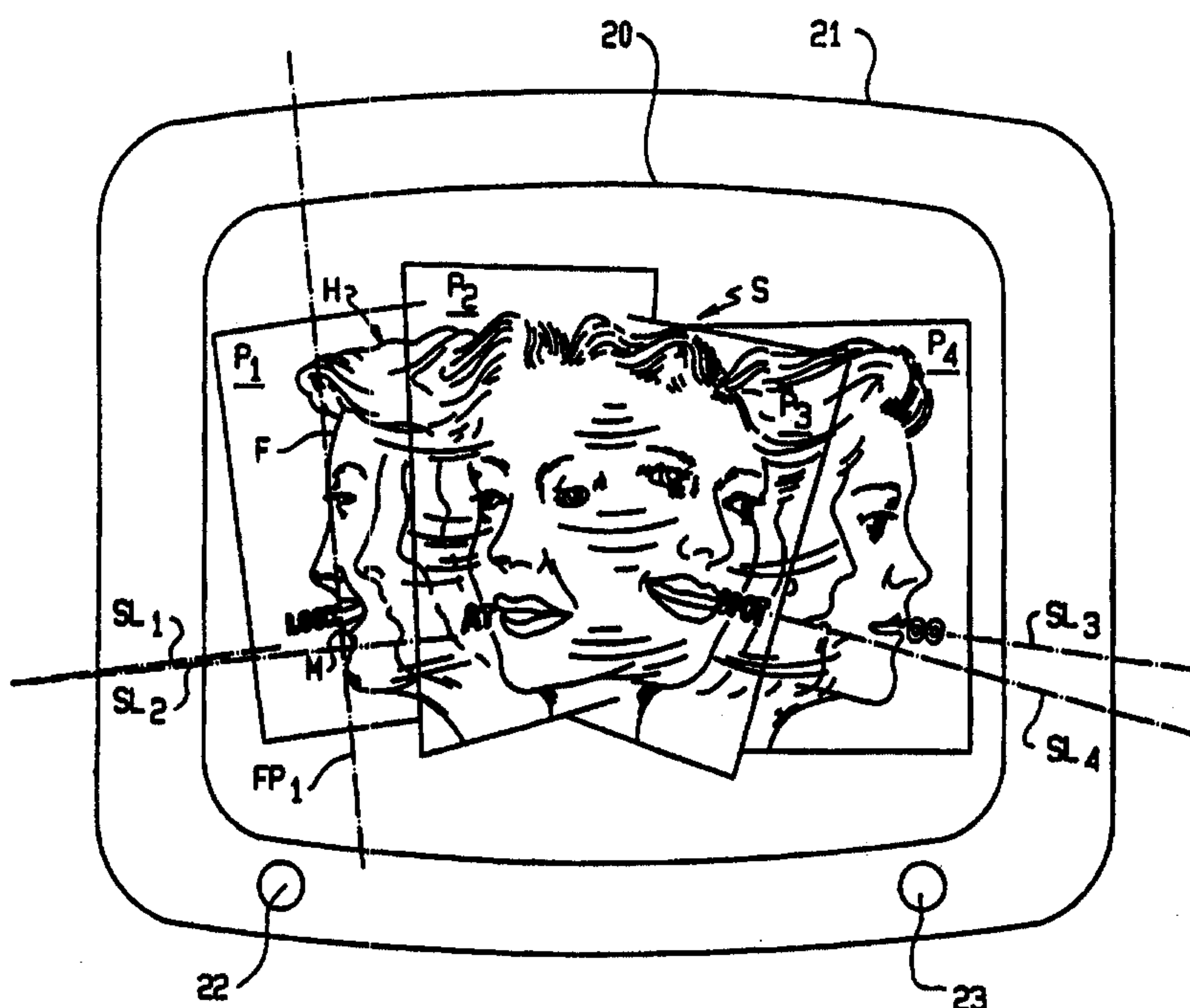
**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : G09B 5/00	A1	(11) International Publication Number: WO 98/49666 (43) International Publication Date: 5 November 1998 (05.11.98)
(21) International Application Number: PCT/US98/08447 (22) International Filing Date: 25 April 1998 (25.04.98) (30) Priority Data: 08/845,863 25 April 1997 (25.04.97) US (71) Applicant: READSPEAK, INC. [US/US]; 135 West 20th Street, New York, NY 10011 (US). (72) Inventor: KIRKSEY, William, E.; 135 West 20th Street, New York, NY 10011 (US). (74) Agents: MCKENNEY, Charles, E. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).		(81) Designated States: AL, AM, AU, AZ, BA, BB, BG, BR, BY, CA, CN, CU, CZ, EE, GE, GH, GW, HU, ID, IL, IS, JP, KG, KP, KR, KZ, LC, LK, LR, LT, LV, MD, MG, MK, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UZ, VN, YU, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, 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, ML, MR, NE, SN, TD, TG). Published <i>With international search report.</i>
(54) Title: METHOD AND SYSTEM FOR MAKING AN AUDIO-VISUAL WORK WITH A SERIES OF VISUAL WORD SYMBOLS COORDINATED WITH ORAL WORD UTTERANCES AND SUCH AUDIO-VISUAL WORK		

(57) Abstract

An audio-visual work and method of its creation which work has writings placed on the pictures of the work so that as each word or other utterance is heard a writing to be associated with the hearing is coordinated with seeing by emphasis or otherwise of the writing such that the future presentation of either the utterance or writing shall evoke the other in the mind of the original viewer-listener. Each word will when appropriate appear in a legible perspective adjacent to the mouth (M) of the utterer (S). The work can be displayed linearly or under computer control of the viewer-listener along with additional educational materials. Selected use of groups of words highlighted or not may be used.



WO 98/49666

PCT/US98/08447

**METHOD AND SYSTEM FOR MAKING AN AUDIO-VISUAL WORK
WITH A SERIES OF VISUAL WORD SYMBOLS COORDINATED
WITH ORAL WORD UTTERANCES AND SUCH AUDIO-VISUAL WORK**

This application is a continuation-in-part of U.S.
5 Application Serial No. 08/310,458 filed September 22, 1994,
entitled "An Audio-Visual Work With A Series Of Visual Word
Symbols Coordinated With Oral Word Utterances", now pending;
which application is a continuation-in-part of U. S.
application Serial No. 08/209,168, filed February 24, 1994,
10 now abandoned, which is a continuation-in-part of U.S.
Application Serial No. 08/126,341, filed September 23, 1993,
now abandoned.

Background of the Invention

15 Prior audio-visual presentations have included placement
of subtitles (U.S. Pat. No. 3,199,115 and U.S. Pat. No.
5,097,349) or balloon-type legends (U.S. Pat. No. 1,240,774
and U.S. Pat. No. 2,524,276), all to assist in language
interpretation of oral portions of the presentation.
20 While prior subtitles have from time to time coincided
with the speaking of a single word in a different language,
such occurrences have been haphazard, infrequent, and without
a controlled pattern to accomplish specific association of a
series sounds with a series of writings. Further, location
25 of subtitle words have been remote from the pictorial action.

Prior art flash cards, each displaying a word, have
attempted to teach reading through repetitive enforced and
unnatural exercise. Although having some effect ultimately,
the use of such cards requires longer periods of learning and
30 the in-person presence of a literate tutor whether a mother
or school teacher. Also such cards do not provide the
strength of association that the present invention delivers
by providing referents within a narrative audio-visual medium
that has appeal to the student outside its literacy-teaching
35 component.

U.S. Patent No. 5,241,671 discloses presenting on a
computer screen the text of any article with some words

WO 98/49666

PCT/US98/08447

underlined and some not underlined. When the user selects a word from the text its definition appears in a window on the screen and an audio pronunciation of the word occurs. An audio sound icon may also be displayed.

- 5 Closed-captioned works provide separate areas or adjacent boxes where groupings of words are displayed. Closed-caption systems display groups of words along the bottom of the screen or at other remote locations away from the speakers or actors. Closed-caption words appear
- 10 alongside, below or above the visual pictorial scene with a different background which background is usually white. The display of sign language symbols with audio-visuals to aid the deaf are also shown in separate adjacent boxes. These box display techniques may be intrusive to viewers.
- 15 Tutorial audio-visuals have been broadcast which include instructors facing the camera and speaking words with the corresponding written words being displayed in front of the speaker as spoken. Viewer-listeners tire of such tutorial formats and particularly, younger viewer-listeners lose
- 20 interest in the subject matter being presented.

Summary of the Invention

- Briefly, the present invention comprises an audio-visual work and its method of creation which utilizes the natural
- 25 setting of commonly-viewed works with their usual and common series of pictorial frames or segments presented along with speech and other oral utterances which works have, in addition, a series of writings thereon which are associated with or correspond to the series of utterances as
- 30 sequentially heard by the viewer-listener. We refer to this as "euthetic" (well-placed) captioning. The spoken word and the written word within this context correspond if they are the same word. A spoken word in one language with a written word having the same meaning in another language are
- 35 associated words in this context.

According to some embodiments of the present invention, it is a feature that each writing appears near, on or in

WO 98/49666

PCT/US98/08447

association with the head of the utterer such that the written word, the spoken word and the accompanying facial, labial and head motion expressions may be simultaneously observed by the viewer/listener and such that an impression
5 is created by the proximity to and alignment with the mouth that the word has emerged from the mouth. According to other embodiments, each writing appears near, on or in association with a hand or hands of a person using sign language.

According to other embodiments of the invention, writing in
10 Braille is "displayed" on a separate device in association with the spoken words of an utterer.

The present invention is used with non-tutorial audio-visuals normally created for entertainment, informational, or other purposes which audio-visuals are not literacy purposed.
15 It may be used with such materials whether as an element of new production or as a retrofit to previously produced audio-visuals. The present invention may also be used for newly produced materials that are literacy-teaching purposed and which are designed for the application of the present
20 invention; such newly produced, literacy-purposed materials embodying the present invention will be enabled by the invention to be less boring and less intimidating to the student than the present literacy-purposed audio-visual materials.

25 It is a feature that the audio-visual work of the invention may be linearly presented or integrated through programming and use of a multimedia computer platform to create a work that is interactively operable by the viewer/listener to provide additional instruction.

30 It is a further feature of the present method that it has utility in a societal effort in which sufficient works are literated, the placement of words on audio-visuals as herein disclosed, using basic words in a language and repetitively broadcasting or otherwise exhibiting such works
35 to a population to teach a segment of the population to recognize such words when reading.

WO 98/49666

PCT/US98/08447

Finally, the invention includes the selected use of groups of words appearing near the source or apparent source of the uttering of the words which words may or may not be selectively highlighted or otherwise differentiated to
5 provide association with a word or words heard having the same meaning or which viewed and heard word or words are to be associated.

Brief Description of the Drawings

10 Fig. 1 is prior art;

Fig. 2 is a series of elevational views of a speaker with written words appearing in different planes at the speaker's mouth;

Fig. 3 is a series of elevational views of the speaker
15 with written words appearing, all in the same plane, at the speaker's mouth;

Fig. 4 is a flow chart showing steps and items of equipment for use in the present invention;

Fig. 5 is a further flow chart showing creation of an
20 interaction work including the simultaneous audio-visual utterance/writing of the present invention;

Fig. 6 is a flow chart showing further steps and items of equipment for using the present invention;

Fig. 7 is a flow chart illustrating a method expanding
25 the audio portion of an audio-visual to assist in coordinating sound and writing;

Fig. 8 is a front elevational view of a speaker with a word near his mouth;

Fig. 9 is a partial schematic plan view of Fig. 8 with
30 dialogue planes shown;

Fig. 10 is a perspective view of a television set screen with a speaker in various positions;

Fig. 11 is another perspective view of another speaker;

Figs. 12a-b are flow charts of a method of carrying out
35 euthetic captioning according to the present invention;

WO 98/49666

PCT/US98/08447

Figs. 13a-b are flow charts of another system and method of carrying out euthetic-captioning according to the present invention.

Fig. 14 is a flow chart of another system and method of
5 carrying out euthetic captioning according to the present invention;

Figs. 15a-b are representations of wave form expansion according to one aspect of the present invention;

Fig. 16 is a flow chart of another system and method of
10 carrying out euthetic captioning according to the present invention;

Fig. 17 is a flow chart of the system and method depicted in Fig. 16 showing further detail regarding the computer workstation;

15 Figs. 18a-b are flow charts showing further details regarding the computer workstation depicted in Fig. 17.

Figs. 19a-d are representations of applying euthetic captioning;

Figs. 20a-b are representations of four-quadrant
20 placement achieved with euthetic captioning according to the present invention.

Figs. 21a-b are flow charts depicting intuitive application of euthetic captioning according to the present invention;

25 Fig. 22 is a schematic diagram of a multimedia platform according to the present invention;

Fig. 23 is a flow chart of an interactive capability according to the present invention;

Fig. 24 is a flow chart of the interactive word
30 pronunciation depicted in Fig. 23;

Fig. 25 is a schematic representation of a blackout zone according to the present invention; and

Fig. 26 is a schematic representation of one embodiment of the present invention using sign language.

35 Fig. 27 is a perspective view of a person with the word "WOW" near her mouth;

WO 98/49666

PCT/US98/08447

Fig. 28 is a view similar to Fig. 27 with two words near the mouth;

Fig. 29 is a view similar to Fig. 27 with three words displayed;

5 Fig. 30 is a view similar to Fig. 27 with three words displayed with one word displayed higher than the other two;

Fig. 31 is a view similar to Fig. 27 with three words displayed with the higher word differently positioned with respect to the other displayed words;

10 Fig. 32 is a perspective view of a person's head in four (4) different dialogue planes with words in such planes;

Fig. 33 is a view similar to Fig. 32 in which a group of words appear in each dialogue plane with selected words emphasized;

15 Fig. 34 is a view similar to Fig. 32 in which the words appear in planes parallel to the flat screen plane with selected words emphasized;

Fig. 35 is a view similar to Fig. 33 in which the words appear in a plane parallel to the flat screen plane with
20 selected words emphasized;

Fig. 36 is a view similar to Fig. 32 with no words emphasized where words are differentiated by initial appearance;

Fig. 37 is a view similar to Fig. 36 with the words
25 appearing in flat planes;

Fig. 38 is a view similar to Fig. 33 with no word emphasized; and

Fig. 39 is a view similar to Fig. 35 with no words emphasized.

30

Description of the Preferred Embodiments

Fig. 1 shows a prior art screen 1 carrying a typical audio-visual picture 2 (shaded area) which have a prior art closed-captioned box 3 within the picture 2 having the words
35 "in the house"; a prior art sign language box 4 in the picture and a lower elongated word tracking area 5 in the picture with the words "at eleven". Area 5 carries words

WO 98/49666

PCT/US98/08447

which move in the direction of arrow A. Sounds including dialogue associated with picture 2 in most part appear to emanate from sound source area 6.

Words or other symbols in accordance with the present
5 invention are normally placed on the pictorial portion of the audio-visual within the sound source area 6; however, words may also be superimposed on that portion of the picture 2 where the listener-viewer's attention is directed by his or her interest in the audio-visual, such as where there is
10 action, whether or not the location of such action coincides with the sound source.

The present invention, in one of the preferred embodiments, places words in the frame of reference of the speakers in the audio-visual (i.e. in planes not parallel to
15 the plane of the viewing screen). Since the frame of reference of the viewer is the plane of the viewing screen, words moved from such plane into the actor's world are more readily and meaningfully viewed and appear to the viewer as three-dimensional objects.

20 Referring to Figures 2-3, speaker S of an audiovisual work has a head H and a mouth M from which the written word "look" appears in plane P_1 as such word is spoken. Plane P_1 is approximately perpendicular to a line through the utterer's ears (not shown). Each word preferably appears
25 during the brief period of time in which the word is spoken or uttered; however, the word may appear in addition just before and just after it is spoken [provided its appearance does not interfere with words and sounds spoken previously or subsequently.] The criterion is that as each word is spoken
30 there is provided to the viewer-listener an associated corresponding written word or writing. The present invention provides for the presentation of a meaningful sequence of spoken words (or other utterances) together with a coordinated sequence of written words, numbers or other
35 writings, to accomplish the association of such spoken word or utterance and writing, one at a time, in the mind of the viewer-listener.

WO 98/49666

PCT/US98/08447

A purpose of the coordination of the presentation of a plurality of written words or writings, one at a time, with corresponding spoken words is to provide the viewer-listener with the [opportunity to associate] in a natural setting such
5 sounds and sights for the purpose of remembering that the sound and sight are to be associated such that future presentations of either the sound or the sight shall evoke the other in the viewer-listener. While this purpose is for literacy, foreign language study and education, another
10 advantage of the invention is increased clarity of understanding in that a viewer-listener may receive and understand the word orally or visually or by both stimuli depending on his or her ability, attentiveness or location with respect to the unit displaying the audio-visual work. A
15 second advantage is the translation of foreign sound tracks with heightened understanding provided by location of the written translation at or near the mouth; and a third advantage is to achieve a simultaneous bilingual writing presentation by presenting two writings, one in the utterer's
20 language and the other in a different language and both occurring simultaneous to the utterance. Where utterances may be delivered in a series so rapid that visual coordination with writings is not practical, that portion of the audio/visual medium so affected may be digitally expanded
25 as to sound and expanded visually by either digital or analogue means [so as to enable comprehensible association.]

Bigrams, trigrams, or quadragrams (two, three, or even four word sequences) may be displayed simultaneously where the goal is comprehension by the deaf or non-speakers of the
30 language of the soundtrack of the audiovisual work and single word presentation is too fast; in both cases, the intention is that the captions will be offered in a language the viewer already understands. In that case, the number of words should be the smallest number of words that will still allow
35 an adequate reading comprehension time window for the phrase in question. This approach is a replacement for closed-captions or foreign film subtitles where the goal is limited

WO 98/49666

PCT/US98/08447

to comprehension of the narrative or entertaining program as opposed to associations with utterances.

According to another embodiment of the present invention, one or more words are positioned in association
5 with the hand or hands of a person speaking in a sign language, such that there is a correspondence between such words and a single sign language element. In this way, the viewer-listener is provided with the opportunity to associate in a natural setting such words for the purpose of
10 remembering that the words are associated with that sign language element.

According to yet another embodiment of the invention, words may be placed on an audiovisual work so that they are visible only to a viewer who uses a special reading device.
15 This is analogous to three-dimensional presentations that are visible only when the viewer wears a special type of eyeglasses. Indeed, a special type of eyeglasses is the preferred method for carrying out this embodiment.

The words of the present invention are displayed as an
20 integral part of and superimposed on the pictorial scene of the work. The pictorial scenes include components such as human figures, furniture, sky, a background citiscape and so forth. The words may be superimposed on one or more pictorial components and by consequence prevent viewing of a
25 portion of the pictorial component or prevent partial viewing of a portion of the pictorial component where the written word is translucent or semi-transparent or the word is composed of wire-framed letters.

Since the presentation of more than one written word to
30 the viewer-listener at one time while the words are being spoken makes it difficult if not impossible to correctly associate the right sound with its corresponding written word, it is important that each sound and its corresponding written word be made available in a manner that makes it easy
35 for the listener-viewer to associate the two elements. To avoid distraction and confusion, each spoken word should be accompanied by its sole written associate with the possible

WO 98/49666

PCT/US98/08447

exception of an added pictogram of such word, sign language representation of such word, or a foreign translation of such word. Such written word or words may be displayed before, during and after the word is spoken, provided that such display does not take place while the preceding word or succeeding word is spoken.

More than one word or symbol may appear during the utterance provided each word and symbol is to be associated with the utterance. For example, if the word "thank-you" is spoken, the word "thank-you" and the word "merci" may simultaneously appear.

During an audio-visual presentation there are speaking periods of time in which words are being spoken and non-speaking periods in between. In the one-word-at-a-time procedure of the present invention the written word appears only during the period comprising (1) the non-speaking period following the speaking-of the prior word (2) the speaking of the word and (3) the non-speaking period following the speaking of the word before the next word is spoken.

By presenting alphabet-based words in a pictorial setting such words, to the mind of the non-literate student, are logograms to be memorized employing that portion of the brain which records whole, visual images, much as a film receives light to create a photograph. The inventive segmental presentation of the alphabet-word in simultaneous accompaniment with either a spoken or pictogrammic referent, or both, creates a recoverable association in the mind of the student between the written word (which is perceived as a logogram although "normally" scripted) and the simultaneously presented referent(s). After some repetition, subsequent presentations of the alphabet-based word (logogram) will recall in the mind of the student the referent(s), i.e., the spoken word. This, of course, defines the act of reading, the teaching of which ability is a purpose of the present invention.

The same process of pairing spoken and written words also teaches, in reverse manner, a student who is literate in

WO 98/49666

PCT/US98/08447

a given language to be able to speak it. In this case, the referent is the written word or logogram and the learning target is the spoken word.

A key to the intensity of the learning, particularly by
5 infants, is that the associations be presented in an environment that is "natural", similar to the environment in which the child learns to speak. The environment in which a child learns to speak, which normally and generally does not include formal speaking lessons, is the same type of
10 environment the present invention delivers audio-visually. In the preferred linear embodiment of this, invention the audio-viewer is provided with an environment of a story or other presentation whose primary purpose is not the teaching of literacy. When one learns to talk, one is exposed to
15 visual images or actions, respectively demonstrated or implied by agencies (such as parents) in the learner's environment, which serve as referents that will achieve association with parallel utterances. The environment of the present invention is one where visual images or actions,
20 respectively demonstrated or implied by agencies (such as parents) in the learner's environment (i.e., a child's), serve as referents that will achieve association with parallel utterances. Such environment includes meaningfully seriatim utterances, inasmuch as agencies in a learner's
25 environment, do not as a rule make random utterances. Such a natural language learning situation is presented in the typical motion picture wherein natural communication situations are depicted and wherein repetitive audience exposure to the same word, through natural recurrences during
30 the film, takes place. The natural environment and the motion picture emulation both provide associations between actions and objects and their corresponding descriptive utterances; the present invention extends the association opportunity to the written word in the audiovisual emulation
35 of the natural environment.

The present method is able to teach reading by presenting to the student whole words as distinguished from

WO 98/49666

PCT/US98/08447

syllables or letters of a word. Viewing and remembering a whole word is akin to viewing and learning a symbol, such as a picture of a cat or a Chinese language character, in that such whole word is, it is believed, processed by the human
5 brain in the same way. Viewing each word as a whole (or sight reading) provides a teaching based on developing associations in the mind that are visually memorized or imprinted and recovered through association rather than through human brain analysis which is required for
10 alphabet-based, syllabic, or phonetic reading.

Where two writings, i.e. one in written form and the other in pictorial form, are caused to be displayed corresponding to a single word spoken, the two writings may merge into or out of one another to indicate that the two are
15 associated or even the same. For example, as a person in a video speaks the word "cat", the written word "c-a-t" could mutate into the pictogram of a cat.

Whether the associations created by the present invention are in the context of an audio-visual now existing
20 or to be created, the associations created by the present invention occur in normal, natural pictorial settings. As examples, such associations could occur in photoplay scenes where a detective and a suspect converse; in videos where a performer sings or in TV newscasts where a weatherman speaks
25 and points to a map. In all the cases just cited, the purpose does not necessarily involve literacy.

The present invention is also applicable to teaching lip reading where as the utterance is made and as the writing is displayed the lip movement is simultaneously made observable
30 as part of the visual portion of the work.

One of the advantages of positioning words at or near the contextual source within the area of the displayed picture is to make it easier for the viewer to see the word as he or she hears the word while maintaining focus on the
35 action in the work as it takes place. Although the eye can see peripherally words positioned at the fringe edges of the viewing screen or even outside the pictured area, it can only

WO 98/49666

PCT/US98/08447

read them with difficulty while still maintaining a meaningful focus on the action elements of the audio-visual work. It is for this reason, among others, that the present invention is superior to closed-captioning. Closed-

5 captioning also presents more than one word at a time, which prevents the association of one word with one sound.

Furthermore, the present invention presents the words in dimensional relation to the speaker which reduces obtrusion and minimizes screen area occupied by the written word.

10 When two people are conversing whether facing one another or not, a single plane between the two people may serve as the plane upon which written words will be displayed. This technique can also be used when one of the speakers is off-camera where the audience is aware of the
15 relative position of the off-camera speaker.

The color, shape and other characteristics of the letters of each written word are designed to be unobtrusive. For example, if the background pictorial component upon which the word is superimposed is a dark blue, the letters of the
20 word may be a light blue or other shade of blue. Also, a written word may be rendered translucently or semi-transparently such that it permits a partial continued viewing of background visuals. Also, a word may be color, font, or otherwise coded to its source.

25 Turning again to Figs. 2-3, as the speaker's (S) head (H) turns, plane P_1 , which is approximately perpendicular to a line through the speaker's ears, moves to three (3) additional positions P_2 - P_4 . As the word "AT" is spoken it appears in plane P_2 in perspective; as the word "SPOT" is
30 spoken it appears in plane P_3 also in perspective and finally as "GO", is spoken it appears in plane P_4 . Each word is located at or near or even on the head and, preferably at or near the mouth (M) of the utterer as it is spoken. Note that as the speaker's (S) head (H) has turned it has also tilted
35 to raise the chin (see plane P_4). Writing orientation

WO 98/49666

PCT/US98/08447

preferably reflects head orientation side-to-side and up-and-down.

In Figure 3, all spoken words appear in planes PP which lie in or are parallel to the screen upon which the audio-
5 visual is presented.

In Figure 4, the apparatus for creating the audio-visual work is described including an operator station; a video text generator to generate the writing desired (such as the word "look") ; audio-visual work input means for providing a work
10 that has had no writings yet placed on it; a digital optical manipulator providing means for combining the text and such audio-visual work to provide the utterance/writing coordination of the present invention in proper plane orientation. This manipulation creates an inventive audio-
15 visual work in which such coordination occurs throughout the work and can be viewed and listened to without interruption in its presentation which embodiment is a linear embodiment of the present invention.

Groups of letters are affixed, imprinted, superimposed
20 or otherwise located on that portion of the picture that is most likely to be viewed as the word is spoken. When the head of the utterer is visible, the location shall generally be at or near the mouth so as to suggest that the word has emerged from the mouth. This sequence is continued for all
25 or a substantial number of utterances for the entire work or, if desired, for a segment of the work. Letters may be of any size, font, or color. In one preferred embodiment, size, font, color, or any other graphic attribute are chosen so as to reflect background colors and the emotional and intentive
30 content of each utterance. As to background, each written word shall be by default translucent, semi-transparent, wire-framed, or in a color that is a shade of the background color, sufficiently differentiated from the background color so as to achieve visibility without leaving a retinal halo or
35 ghost image once the word is gone. As to emotion, intent, or meaning, angry words, for example, will have a red blush with a sharp-edged typeface while lullaby lyrics will be pastel

WO 98/49666

PCT/US98/08447

tinted with a soft, cursive typeface. Emotionally neutral words will be presented in the default color. The purpose of the graphic attributes is to provide the viewer listener with a dynamic graphic parallel to the nuances of the utterances rendered through the variables of volume, tone, pitch, or other vocal attribute and to thereby enhance the goal of an association that is recoverable in the future by the mind.

Natural communication situations are prevalent in audio-visual works. Such situations include a detective interrogating a suspect as referred to above. Placing words on scenes including natural communication situations provides a vehicle for creating the association of sound and writing desired while the viewer-listener remains attentive to the natural communication of the work.

Turning next to Fig. 5, the linear embodiment of the invention is used to create an interactive embodiment by creating a computer program permitting the viewer/listener to stop the audio-visual presentation to bring up for viewing on the screen on which the audio-visual is being presented a menu for providing by selection, word definitions, syntax and sentence context usage or other information. The interactive work is presented by operation of the viewer/listener using a programmable educational apparatus for using such program to display the work, stopping the work to view a selected writing and to obtain additional information relating to such writing.

Turning to Fig. 6, audio-visual works are created by a computer graphic designer at his or her work station where the video signal of the work (in analogue or digital form) is presented on a screen to the designer. In working with frames (pictorial sequences of 1/30th of a second), the designer creates a computer graphic or text (i.e. a word) and superposes it on the video signal of the frame or frames depending on the length of time the speaking of the word takes. The length of time it takes to speak a word varies with a large number of words in everyday English (or other language) conversation taking between 1/60th and 1/2 of a

WO 98/49666

PCT/US98/08447

second. By employing animation and using paint box software additional characters may be given to the font of letters in the word and the orientation of the word in a selected plane.

Fig. 7 illustrates the method extending the time a word
 5 is spoken in an audio-visual for the purpose of providing longer presentation of the associated written word. This extension or spreading out of the time a word is heard is accomplished by digitizing the sound of the word on a hard disk as a wave form and then reconfiguring the wave form.
 10 Such a technique does not distort the pitch or the tone.

Head (H') of Fig. 8 is facing to the viewer's right as indicated by dashed source line (SL). Line (SL) lies in speaker reference dialogue plane (P₅) (not shown). Vertical viewer reference plane (A) is viewed by the viewer as a line.
 15 This plane remains fixed. Line (SL) goes through word "WOW" like a barbecue skewer.

The distance the beginning of the word (WOW) is positioned from the head (H') of a speaker is preferably within a distance (d₂) which is twice the width (d₁) of the
 20 speaker's face (F) having nose (N) (see Fig. 8). This positioning of the word (WOW) in the range of 2 d₁ provides good results for scenes where the speaker's head is in a close-up position. Where the head is distant as in a long shot, the word may be larger than the head but still adjacent
 25 to head (H') or shifted to an object of viewer-interest and, in such instance, distance (d₂) may be 3 or 4 times distance (d₁).

Fig. 9 is a schematic plan view of Fig. 8 showing dialogue plane (P₅), plane A (the 180° viewer reference plane)
 30 and B,, the 90° plane. Dialogue plane (P₅) which has source line (SL) therein includes the word "WOW" which appears in such orientation. Words appearing in other dialogue planes (P₆) and (P₇) Which are 25° from viewer plane (A), the 180° viewer reference plane, are readable but since words placed
 35 in dialogue planes closer to viewer reference plane (A) (the viewer's principle plane of vision) are difficult to read

WO 98/49666

PCT/US98/08447

such positioning (in this "blockout area") is rarely used in the practice of this invention.

Fig. 10 shows television screen 20 of set 21 with control knobs 22, 23. The speaker's head/face position is shown in multiple views as it was shown in Fig. 2. The view to the left of screen 20 shows head (H), face (F), dialogue plane (P₁) with source line (SL₁) in such plane. Face plane (FP₁) is perpendicular to the dialogue plane (P₁). Source line (SL₁) is perpendicular to face plane (FP₁). Face planes generally lie in planes perpendicular to the horizontal when the speaker is in or standing or sitting position. Source line (SL₁) bisects linearly the word "look". Other source lines (SL₂), (SL₃) and (SL₄) are shown lying in their respective dialogue planes (P₂), (P₃) and (P₄) each of which lines bisects linearly its respective word.

Finally, Fig. 11 shows a speaker (S₂) with head (H₃) face (F₃) and mouth (M). Face plane (FP₂) is perpendicular to a dialogue plane (not shown in this figure). Source line SL₃ which lies in the dialogue plane (not shown) bisects the word "fast". Since head (H₃) may move in any orientation as speaker (S₂) reclines or turns her back to the viewer, words on source line (SL₃) as spoken by head (H₃) in such orientation are in each instance placed in the dialogue plane except where the dialogue plane's orientation is such that the word as placed lacks legibility to the viewer. For example where speaker (S₂) is in a standing position and facing away from the viewer, the word "fast" if placed in the dialogue plane would be at an angle to the viewer where the word "fast" would be illegible. To avoid such illegibility the word is placed in a plane as close to the dialogue plane as possible where the word "fast" is legible. In such a case the word "fast" would be shown in a perspective orientation in such selected plane to give the impression that the word was going away from head (H₃).

Where time permits, the word "fast" may originally appear in a position obscuring a portion of the mouth (M) and

WO 98/49666

PCT/US98/08447

then be moved quickly along the line (SL₅) of the dialogue plane. Alternatively, for example, if the word is to appear on the screen for .024 thousandths of a second, the word may appear for .008 thousandths of a second partially on mouth
5 (M) and then move along line (SL₃) for .008 thousandths of a second and finally stop on the line for another .008 thousandths of a second before disappearing.

The purpose of placing words on a source line (SL) and in planes in perspective as set out herein is to cause the
10 word to appear if it came out of a mouth and thereafter appeared as an object in the three-dimensional space of the audio-visual scene. As an object, the written word is subject to the same physical laws that any other object is subject to.

15 Thus, if someone walks in front of a speaker in an audiovisual work using the present invention, the speaker's speech may be muffled and view of his written word may be momentarily blocked partially or wholly. The purpose of this aspect of the invention is to make the words appear to be
20 real objects, a concept very acceptable to young minds in particular who will find the words "user friendly" rather than abstract.

Words are positioned to appear in perspective with the letters of the words increasing or decreasing in size (see
25 Fig. 8 where the "w" to the left is smaller than the "o" which in turn is smaller than the "w" to its right). Words in perspective appear to have direction including the appearance of moving in such direction. A word in perspective near a speaker's mouth appears to be coming from
30 the mouth. Words are placed as close to the mouth as possible without interfering with those facial expressions of the speaker which are part of the communication.

Not all words spoken during a work need have a corresponding written word displayed since selected periods
35 of running of the work may offer special difficulties in literation or for other reasons may not require literation.

WO 98/49666

PCT/US98/08447

The preferred use of the invention is in emplacement of the words or other alpha numerical symbols or other writings on tapes, films, computer diskettes, CD ROMS or other media in a meaningful sequence which provides association with the
5 oral component of the tape or film or CD ROM or computer diskette in the manner described above. Such sequencing may continue throughout the film or tape from beginning to end. Audio-visual works of the present invention have preferably entertaining or otherwise contextually meaningful subject
10 matter and content. The learning by the viewer/listener occurs without specific effort on his or her part as he or she enjoys the entertaining or other subject matter.

The present invention creates within a pictorial area of the work an impression of the spoken word as if it were
15 visible in that each word, as viewed, has dimension, color, font, motion and other characteristics. The dimension of the word is the orientation of the word in the plane of the display screen or in a plane at an angle to such plane. Words in such orientation are three-dimensional as are other
20 components of the picture.

Writings may include letters, words, pictures or other symbols.

According to another embodiment of the present invention, the writings are displayed in Braille, preferably
25 on a separate device that a person (e.g. a sight-impaired person) can use while listening to an audio program. Analogous to other embodiments, a one at a time correspondence is established between the Braille writings and the spoken utterances, such that the user is provided
30 with an opportunity to associate in a natural setting such writings for the purpose of remembering that the writings are associated with those utterances.

Example

35 An entertaining video game is employed in which an inventory of pictogrammic (literal drawings) referents are available to the player. The pictograms will be cursor

WO 98/49666

PCT/US98/08447

draggable. One mouse click on any referent will result in the referent fading into ("morphing") its written word equivalent (logogram) while a voice-over or talking head utters the word.

5 A goal of the game is to create a row of pictogrammic referents which creates a meaningful seriatim. Once the player has arranged such a row, a double-click of the mouse will result in the referents morphing into written words (logograms), from left to right, one at a time, and in
10 simultaneous accompaniment with the appropriate spoken referent. Then the meaningful seriatim is repeated aloud, left to right, by the utterer, each word being suddenly "backgrounded" by a referent.

In playing of the game a drag created arrangement of
15 referents that is not meaningfully seriatim will result in no outcome when double-clicking is undertaken and no points are scored.

Nuances of color and font graphics may vary in accordance with the natural flow of the meaningful expression
20 of dialogue. As such, the overall "organic look" of the invention will create a novel, standardized "emotive graphic vocabulary". As examples, the following colors and graphics may be used for the following emotions:

25	<u>Emotion</u>	<u>Color</u>	<u>Graphic</u>
	Happy	white or Pink	
	Twinkle/sparkle		
	Sad	Blue or Black	Gothic/
	Angry	Red	Bold
30	Sexual	Purple	Undulating

Font and color nuances might also be used to associate physical realities, such as found in nature.

35

WO 98/49666

PCT/US98/08447

<u>Physical</u>	<u>Color</u>	<u>Graphic</u>
Cold	Gray/Ice-Blue	Icicle
Hot	Orange/Red	Flame
5 Wet	Milky	Drop

Such associations are based on common sense and/or preexisting studies linking the associative graphic effects of color, texture, etc., on human emotions and learning retention. In addition, the capabilities of the present
 10 graphic computer software including visual phenomena, such as "glowing" and "radiating," can be layered in for additional associative impact.

Euthetic captioning in a narrative context according to the present invention may be accomplished in a number of
 15 ways. Figs. 12a-b show steps for applying euthetic captioning manually.

Figs. 13a-b depict a video direct system and method of applying euthetic captions.

Fig. 14 depicts a system and method that slows down
 20 utterances without loss of pitch or tone and without apparent distortion. Figs. 15a-b are depictions of a normal and expanded waveform, respectively, of the word "future" as expanded by the system and method depicted in Fig. 14. The waveform of Fig. 15b has the same pitch as the waveform of
 25 Fig. 15a because the amplitude of the waveform is kept constant while waveform is expanded.

Another embodiment of the invention is useful when the waveform is expanded by some fractional multiplier, as opposed to a whole number multiplier. For example, when it
 30 is desired to increase the length of a waveform by one-half (a 50% increase), as opposed to doubling the length (a 100% increase), known methods randomly select which portions of the waveform to expand. According to this aspect of the invention, the random selection of portions of the waveform
 35 is restricted to only vowel portions. This may be accomplished by means known to those of skill in the art.

WO 98/49666

PCT/US98/08447

Fig. 16 depicts a digital system and method of applying euthetic captioning, utilizing known character animation software to position words.

Figs. 17-21 relate to another embodiment of the present invention, which is a system and method for intuitive euthetic captioning. Fig. 17 shows the system and method depicted in Fig. 16, utilizing euthetic captioning software according to this embodiment of the invention. Fig. 18 is a flow diagram showing further detail of the computer workstation used in the system and method depicted in Figs. 16 and 17.

Figs. 19a-d demonstrate details of applying euthetic captioning according to the present invention. Figs. 20a-b depicts the four quadrants of a virtual three-dimensional world that a euthetically captioned word appears to inhabit. Fig. 21 is a flow diagram of a software module for placing a euthetically captioned word in an optimal orientation in any quadrant of the virtual three-dimensional world. The software preferably runs on a computer workstation system. While many input devices known to those of skill in the art may be utilized, preferably the user specifies a quadrant and draws a source line with a mouse, and enters with a keyboard. The computer system running the software module automatically positions the word on the source line, preferably so that the source line runs through the center of the main body of lower case letters (known to typographers as the "x-height"), such as the horizontal bar in the letter "e". once the system and software have placed the word, the source line is deleted.

Fig. 22 is a schematic that depicts a multimedia platform incorporating an interactive multimedia computer workstation for creating interactive euthetically captioned works according to the present invention. Fig. 23 is a flow diagram of software to implement interactive capabilities.

Fig. 24 is a flow diagram of interactive word pronunciation depicted in Fig. 23. According to this aspect of the invention, when a user stops a euthetically-captioned audiovisual work on a particular word, the user may obtain a

WO 98/49666

PCT/US98/08447

pronunciation lesson. Preferably, the user may speak into a microphone connected to a computer that contains voice wave analysis software, which compares the wave form created from the user's pronunciation of the word to a standard wave form
5 for the correct pronunciation stored in a computer file. The computer then provides feedback to the user that either confirms correct pronunciation (for example, as "good enough" or "excellent") or prompts the user to try to pronounce the word again.

10 The other options depicted in Fig. 23 preferably will be presented as a menu of interactive applications that a user may select. For example, the user may select a writing application that will allow the user to mimic a word displayed by typing the word or by writing the word on an
15 electronic tablet that produces output to handwriting recognition software. The interactive system preferably would provide feedback to inform the user whether or not the word had been properly typed or written.

Fig. 25 represents the placement of a word in the
20 frontal "blockout zone" depicted in Fig. 9. A word may optionally be placed in this zone -- on a plane passing through line AA of Fig. 25 -- as one way to make it appear that it is emanating from the speaker's mouth.

One or more words may also be placed in association with
25 the hand or hands of a person using sign language, such that there is a correspondence between such words and a single sign language element. An example of this embodiment of the present invention is depicted in Fig. 26, which shows a speaker on a TV screen and an inset box with a hand of a
30 person doing simultaneous sign language translation. As the speaker says the word "Future" and the sign language interpreter signs that word, "Future" is placed in the inset box in association with the sign language element for that word.

35 Turning to Figs. 27-31, a series of words are shown emanating from the utterer's mouth (M). In Fig. 27, the word "WOW" is in the vicinity of the mouth (M) along speech line

WO 98/49666

PCT/US98/08447

(SL). Speech line (SL) projecting from mouth (M) is in an orientation such that the words appear to have just issued from the mouth as objects. In Fig. 28, a second word "HE'S" has appeared behind the word "WOW" which has moved to provide
5 space for the word "HE'S". The word "WOW" has also been differentiated from the word "HE'S" through emphasis of the word "WOW" and in Fig. 29, a third word "FAST" has appeared and has been differentiated by bold letter emphasis. In Fig. 30, the word "WOW", now emphasized, has moved upward but is
10 still near the head. Fig. 31 shows the word "WOW" remotely positioned with respect to the other words displayed and is above and to the right of the head. All the words of Figs. 27-31 are in the vicinity of the head and the words of Figs. 27-30 are in the vicinity of the mouth.

15 A further embodiment of the invention is shown of Fig. 32 in which words appear successively one new word at a time in sequential scene presentations. The words as they accumulate are bundled near the mouth with some words above other words, all words being displayed adjacent the mouth of
20 the utterer in dialogue planes P_1 - P_4 which planes have differing orientations as the utterer's head moves. Bundling of words is positioning them to occupy areas which have height and width dimensions which are different than dimensions of an area a string of words would occupy which
25 area would include an elongated rectangle where the width is substantially greater than the height. The purpose of bundling including using two or more lines or rows of words is to position all the words near to source or apparent source of the uttered words.

30 Turning back to Fig. 32, the word display in plane P_1 occurs first followed by the displays in planes P_2 - P_4 . In plane P_2 the words "LOOK AT" are in a row just above the utterer's lips. In the subsequent plane P_2 after the utterer has turned her head, the words "LOOK AT" appear in a row
35 similarly located with "LOOK" closer to the lips and "AT" further away from the lips. The word "SPOT" appears in plane P_3 in a lower row and is in bold letters. Planes P_2 - P_4 are

WO 98/49666

PCT/US98/08447

dialogue planes in which each word lies and differentiated from other words and pictorial objects by bold lettering for a period of time during which period the differentiated word is heard by the viewer-listener.

5 Through the use of differentiating one word from other words at the same time the differentiated word is heard by the viewer-listener and association is established between the differentiated word being focused on by the viewer-listener and the word being heard. Differentiation may be
10 accomplished by any means which causes the viewer-listener to distinguish the differentiated word from other words or objects displayed. In this embodiment, all words displayed that are in the vicinity of the mouth and near the head of the utterer. Examples of techniques for differentiating
15 words are causing the word or any part thereof to enlarge to a particular size, or to have a particular shading, color, shape, or brilliance or to cause the word to appear to move actually or to move or otherwise create a visual effect which calls attention to the word where other words are present and
20 to thereby differentiate such word from such other words.

In another embodiment of Fig. 33, an entire group or bundle of words are shown in each dialogue plane P_1 - P_i for a selected period of time and during a portion of such period of time each word differentiated by bold-letter emphasis is
25 spoken. The differentiation and the speaking or uttering occur at the same time. In this embodiment, no words are displayed that are not in the vicinity of the mouth of the utterer. Figs. 34 and 35 illustrate embodiments which are similar to Figs. 32 and 33 except the words appear in planes
30 (PP) parallel to the viewing screen.

In embodiment Fig. 36 the technique of word differentiation is found in its introduction. The word is spoken at the same time the word is first introduced or appears on the screen. When the word "LOOK" first appears on
35 the screen, the word "LOOK" is therefore heard by the viewer-listener to create the desired association. While the word "LOOK" is still on the screen the word "AT" appears and at

WO 98/49666

PCT/US98/08447

that time the word "AT" is heard. In the third sequence, the viewer-listener sees "LOOK AT" near the utterer's mouth and as the listener-viewer hears the word "SPOT" the word "SPOT" is first seen. Fig. 37 illustrates the technique of
5 Fig. 36 except the words appear in flat planes.

The one or more words of Figs. 27-37 embodiments preferably appear in the viewing vicinity of the mouth of the utterer as well as in a frame portion near the head. The words also preferably appear to the viewer-listener in an
10 order relating to their meaning as a term, phrase, or group of words. The words are uttered and heard in an ordered way during the period of the viewer-listener views the words near the head.

The final embodiment of the invention employs the
15 technique of group display of words on the screen (see in Figs. 38-39). This embodiment has two specific purposes. One purpose is to provide to the viewer-listener the gist of what is being spoken in a language not known to the viewer-listener. The second purpose is to provide the gist of words
20 being spoken but not intelligible to the viewer-listener due to hearing or other impairments or due to inaudibility of the sound. The bundled words may be placed in dialogue planes or planes parallel to the first screen plane. In this embodiment no words need be emphasized.

25

30

35

WO 98/49666

PCT/US98/08447

I CLAIM:

1. A teaching audio-visual work including pictorial scenes with natural communication situations for presentation to a viewer-listener comprising

5

(a) at least one pictorial scene comprising frames capable of having its visual portion displayed on a screen and its audio portion heard;

10

(b) a series of word utterances by at least one utterer with each utterer having a head in successive frames of such scene presentation;

15

(c) a series of visual word symbols, one or more word symbols being displayed in a frame portion near the head of the utterer for a display time period with one or more such word symbols displayed in the near frame portion being differentiated for a differentiation time period, each differentiated word symbol

20

(i) corresponding to and associable with each such word utterance,

25

(ii) capable of teaching the viewer-listener the visual appearance of the uttered word;

30

(iii) appearing with such differentiation in the near frame portion so that the viewer-listener has an opportunity to be able to make such association, and

35

(iv) appearing differentiated in such frames of the pictorial scenes, at the same time the differentiated

WO 98/49666

PCT/US98/08447

5

word utterance is heard so that each such utterance and each such corresponding differentiated word symbol are associable at the same time and for such length of time by the viewer-listener,

whereby the viewer-listener does associate each such utterance with each such word differentiated symbol to learn
10 the visual appearance of the uttered word and the sound of the word symbol.

2. The audio-visual work of claim 1 in which the one or more word symbols appear to exit sequentially from the
15 mouth with selected exiting words being differentiated at the time the selected word is uttered.

3. The audio-visual work of claim 1 in which each word symbol is differentiated after first appearing at the mouth.
20

4. The audio-visual work of claim 1 in which differentiation is by emphasizing the word which word is de-emphasized at the end of the differentiation time period and a second displayed word is emphasized for a second
25 emphasisization time period.

5. The audio-visual work of claim 1 in which word symbols exit the mouth from the utterer and while moving away from the mouth remain near the head and thereafter move to a
30 portion of the screen remote from the head.

6. The audio-visual work of claim 1 in which the utterer has ears and in which the word symbols appear in a dialogue plane passing substantially perpendicular to a line through
35 the utterer's ears.

WO 98/49666

PCT/US98/08447

7. The audio-visual work of claim 1 in which the presentation includes non-tutorial scenes with such word symbols superimposed on such scenes.

5 8. The audio-visual work of claim 1 in which the word or words appear in a dialogue plane and such dialogue plane is at least sixty (60) degrees from such flat screen plane.

9. The audio-visual work of claim 1 in which at least
10 one word symbol is a logogram.

10. The audio-visual work of claim 1 in which at least one word symbol is a pictogram.

15 11. The audio-visual work of claim 1 in which the pictorial scenes are presented without interruption in a linear manner.

12. The audio-visual work of claim 1 in which the
20 pictorial scenes are presented in an interrupted non-linear manner by use of a computer operable by the viewer-listener.

13. The audio-visual of claim 1 in which a portion of the audio-visual work is expanded to extend the length of
25 time the one or more word symbols are seen.

14. The audio-visual work of claim 8 in which a dialogue plane is positioned through the head of an utterer and in which at least one word symbol appears in such
30 dialogue plane or close to such dialogue plane.

15. The audio-visual work of claim 1 in which an utterer's head includes a face and lips and in which facial, labial and head motion associated with the word utterances
35 are included in such frames.

WO 98/49666

PCT/US98/08447

16. The audio-visual work of claim 1 in which the differentiated word symbol appears for a selected time and during such time the word symbol is obscured by an object in the pictorial scene as frames change.

5

17. The audio-visual work of claim 1 in which the word symbols are shaped and angled to appear to be three dimensional.

10

18. The audio-visual work of claim 1 in which the series of utterances are a series of words in sentence form having meaning.

19. An audio-visual work for presentation to a viewer
15 comprising

a) plurality of scenes;

20

b) at least one scene comprising frames capable of having their visual portion displayed on a screen and their audio portion heard;

25

c) one or more word utterances by at least one utterer with each utterer having a head appearing in successive frames of such scene presentation which word utterances are heard by the viewer-listener for an utterance time period;

30

d) one or more word symbols appearing in the vicinity of the utterer head during the utterance time period which one or more word symbols;

35

i) correspond to and are associable with said one or more word utterances;

WO 98/49666

PCT/US98/08447

ii) are positioned relative to the utterer's mouth;

iii) convey the meaning of the word utterances,

5

whereby the viewer may associate the uttered word or words with viewed word or words and whereby the viewed word or words provide the meaning to the viewer of the uttered word
10 or words.

15

20. The audio-visual work of claim 19 in which the uttered word or words are a group of words having meaning and in which the word or words displayed have the same meaning.

20

21. The audio-visual work of claim 19 in which the utterer has ears and in which the word symbols appear in a dialogue plane passing substantially perpendicular to a line through the utterer's ears.

25

22. The audio-visual work of claim 19 in which the presentation includes non-tutorial scenes with such word symbols superimposed on such scenes.

23. The audio-visual work of claim 19 having a dialogue plane and in which such dialogue plane is at least sixty (60) degrees from such flat screen plane.

30 24. The audio-visual work of claim 19 in which at least one word symbol is a logogram.

25. The audio-visual work of claim 19 in which at least one word symbol is a pictogram.

35 26. The audio-visual work of claim 19 in which the pictorial scenes are presented without interruption in a linear manner.

WO 98/49666

PCT/US98/08447

27. The audio-visual work of claim 19 in which the pictorial scenes are presented in an interrupted non-linear manner by use of a computer operable by the viewer-listener.

5 28. The audio-visual of claim 19 in which a portion of the audio-visual work is expanded to extend the length of time the one or more word symbols are seen.

29. The audio-visual work of claim 19 in which a
10 dialogue plane is positioned through the head of an utterer and in which at least one word symbol appears in such dialogue plane or close to such dialogue plane.

30. The audio-visual work of claim 19 in which an
15 utterer's head includes a face and lips and in which facial, labial and head motion associated with the word utterances are included in such frames.

31. The audio-visual work of claim 19 in which the word
20 symbols appear for a selected time and during such time one or more of the word symbols is obscured by an object in the pictorial scene as frames change.

32. The audio-visual work of claim 18 in which the
25 word symbols are shaped and angled to appear to be three dimensional.

33. A method of causing a population group comprising members to become aware of a group of written words of the
30 language spoken in a territory comprising

(a) causing a plurality of audio-visual works to
be created each of which works include a
plurality of pictorial segments including a
series of utterances with each segment;
35

WO 98/49666

PCT/US98/08447

- 5 (b) causing to appear on such segment one or more written words corresponding to such utterances in such a way that each utterance and each written word are associated by differentiation or otherwise with the corresponding utterance;
- 10 (c) providing at least one segment in said works for each word in the group of written words; and
- (d) broadcasting and otherwise introducing such works into the territory to an extent and for a period of time
- 15 whereby the population group becomes aware of such written words of the language.

34. A method of teaching a student comprising

- 20 (a) creating one or more audio-visual works including natural communication situations, each of which works includes presentation during such situations of a plurality of utterances simultaneously with corresponding
- 25 writings, each audio-visual including an area to which the student's attention is naturally directed, which area includes (i) a first portion from which utterances either appear to emanate or to which the student's eye is drawn
- 30 by the invitation of the meaning of the utterances and (ii) a second portion displaying one or more writings simultaneously with each utterance such that an association between the utterance and a differentiated
- 35 corresponding writing will occur in the mind of the student;

WO 98/49666

PCT/US98/08447

(b) making such works available to the student;
and

5 (c) allowing such student to select arrangements
for viewing and listening over time until each
of the plurality of utterances has been heard
by the student together with the viewing of
their corresponding writings a sufficient
10 number of times to assist in learning that
certain utterances correspond to such certain
writings.

35. The method of claim 34 in which the first and
second portions overlap.

15

36. An audio-visual work including pictorial scenes
with natural communication situations for presentation to a
viewer-listener comprising

20 (a) a series of utterances by a human or other
utterers in such scene presentations;

(b) a series of writings associated with such
series of utterances with a writing
25 particularized by initial appearance or
emphasis being briefly located within the
pictorial scenes which writing corresponds
with the utterance heard so that each
utterance and the writing are associated in
30 the mind of the viewer-listener.

37. The method of claim 36 having the additional step
of placing one or more word symbols in a dialogue plane.

35 38. The method of claim 36 in which the series of
utterances are a series of words in sentence form having
meaning.

WO 98/49666

PCT/US98/08447

39. The method of claim 36 having the further step of providing the utterer with ears and causing word symbols to appear in a dialogue plane passing substantially perpendicular to a line through the utterer's ears.

5

40. The method of claim 36 having in addition the step of presenting non-tutorial scenes in the pictorial scenes with such one or more word symbols superimposed on such non-tutorial scenes.

10

41. The method of claim 36 in which the audio-visual is viewed on a flat screen plane and having the step of positioning such dialogue plane at least sixty (60) degrees from such flat screen plane.

15

42. The method of claim 36 in which at least one word symbol is a logogram.

43. The method of claim 36 in which at least one word
20 symbol is a pictogram.

44. The method of claim 36 having the step of presenting pictorial scenes without interruption in linear manner.

25

45. The method of claim 36 having the step of presenting pictorial scenes in an interrupted non-linear manner by use of a computer operable by the viewer-listener.

30 46. The method of claim 36 having the step of expanding a portion of the audio-visual work to extend the length of time the first area word symbol is seen.

47. The method of claim 36 in which a dialogue plane is
35 positioned through the head of an utterer and in which at least one word symbol appears in such dialogue plane or close too such dialogue plane.

WO 98/49666

PCT/US98/08447

48. The method of claim 36 in which an utterer's head includes a face and lips and in which facial, labial and head motion associated with the word utterances are included in such frames.

5

49. The method of claim 36 in which one or more word symbols appear for a selected time and during such time the one or more word symbols are obscured by an object in the pictorial scene as frames change.

10

50. The method of claim 36 in which the word symbols are shaped and angled to appear in three dimensions.

51. The method of claim 36 in which the series of
15 utterances are a series of words in sentence form having meaning.

52. A method of positioning a series of writings on pictures, frames or segments of an audio-visual presentation
20 which pictures, frames, or other segments have a human or inanimate utterer thereon at the time such utterance is made comprising

- 25 (a) selecting an operator controlled unit of equipment including a video text means;
- (b) causing such video text means to display a plurality of words, each word having a variety of sizes, shapes and orientation for
30 retrieval;
- (c) positioning each of a series of pictures for viewing by the operator of the unit;
- 35 (d) operating the video text means to select the displayed words;

WO 98/49666

PCT/US98/08447

(e) conveying the selected words on to a selected picture, frame or segment for permanent location thereon; and

5 (f) causing selected words to be emphasized

so that the word appears on a segment in both local and temporal association with the utterance of the word.

10 53. The method of claim 52 in which the words are placed near the head of the utterer.

54. A computer-based data processing system for euthetic captioning of a plurality of pictorial segments
15 including utterances with each segment, comprising:

(a) computer processor means for processing data;

(b) storage means for storing data;

20

(c) means for superimposing on data representing the plurality of pictorial segments including utterances with each segment, data representing written one or more words which
25 one or more words are emphasized to correspond to such utterances in such a way that each utterance and each written word are associated.

55. A system as claimed in claim 54, wherein the means
30 for superimposing comprises:

(a) means for inputting a first analog video signal;

35 (b) means for converting the analog video signal to video digital data;

WO 98/49666

PCT/US98/08447

- (c) means for displaying the video digital data one frame at a time;
- 5 (d) means for inputting from a user word data and quadrant data;
- 10 (e) means for incorporating the word data, in accordance with the quadrant data and other predetermined criteria, in the video digital data;
- (f) means for converting the video digital data to a second analog video signal;
- 15 (g) means for outputting the second analog video signal.

56. A system as claimed in claim 54, wherein the means for inputting from a user word data and quadrant data
20 comprises:

- (a) means for allowing the user to select a quadrant;
- 25 (b) means for allowing the user to draw a source line;
- (c) means for calculating an angle for the source line;
- 30 (d) means for allowing the user to input the word data; and
- (e) means for automatically positioning the word
35 along the source line.

WO 98/49666

PCT/US98/08447

57. A computer-based system for allowing a user to interactively operate a euthetically captioned audiovisual work, comprising:

- 5 (a) means for displaying the audio-visual work;
- (b) means for pausing the display of the audio-visual work;
- 10 (c) means for allowing the user to specify a word;
- (d) means for processing further data regarding the word specified by the user.

15 58. A system as claimed in claim 57, wherein the means for allowing the user to specify a word allows the user to specify a word that is displayed when the display of the audio-visual work is paused.

20 59. A system as claimed in claim 57, wherein the means for processing further data regarding the word comprises;

- (a) means for storing standard data representing proper pronunciation for a plurality of words;
- 25 (b) means for inputting from the user voice data representing the user's pronunciation of the word;
- 30 (c) means for comparing the voice data to the standard data for the word; and
- (d) means, responsive to the means for comparing the voice data, for indicating to the user
- 35 whether or not the user's pronunciation of the word was correct.

WO 98/49666

PCT/US98/08447

60. A method of providing the meaning of one or more words heard by a viewer-listener in an audio-visual work during a hearing period of time but not comprehended by the viewer-listener comprising

5

1) presenting a plurality of scenes in such audio-visual work having in each scene an utterer with a head and mouth;

10

2) causing one or more word symbols to appear in the vicinity of the said utter's mouth during all or part of the hearing period time which one or more words convey the meaning of the one or more words heard.

15

61. The method of claim 60 having the additional step of placing one or more word symbols in a dialogue plane.

62. The method of claim 60 in which the series of
20 utterances are a series of words in sentence form having meaning.

63. The method of claim 60 having the further step of providing the utterer with ears and causing word symbols to
25 appear in a dialogue plane passing substantially perpendicular to a line through the utter's ears.

64. The method of claim 60 having in addition the step of presenting non-tutorial scenes in the pictorial scenes
30 with such one or more word symbols superimposed on such non-tutorial scenes.

65. The method of claim 60 in which the audio-visual is viewed on a flat screen plane and having the step of
35 positioning such dialogue plane at least sixty (60) degrees from such flat screen plane.

WO 98/49666

PCT/US98/08447

66. The method of claim 60 in which at least one word symbol is a logogram.

67. The method of claim 60 in which at least one word symbol is a pictogram.

68. The method of claim 60 having the step of presenting pictorial scenes without interruption in linear manner.

10

69. The method of claim 60 having the step of presenting pictorial scenes in an interrupted non-linear manner by use of a computer operable by the viewer-listener.

15

70. The method of claim 60 having the step of expanding a portion of the audio-visual work to extend the length of time the first area word symbol is seen.

71. The method of claim 60 in which a dialogue plane is positioned through the head of an utterer and in which at least one word symbol appears in such dialogue plane or close to such dialogue plane.

72. The method of claim 60 in which an utterer's head includes a face and lips and in which facial, labial and head motion associated with the word utterances are included in such frames.

73. The method of claim 60 in which one or more word symbols appear for a selected time and during such time the one or more word symbols are obscured by an object in the pictorial scene as frames change.

74. The method of claim 60 in which the word symbols are shaped and angled to appear in three dimensions.

WO 98/49666

PCT/US98/08447

75. The method of claim 60 in which the series of utterances are a series of words in sentence form having meaning.

5

10

15

20

25

30

35

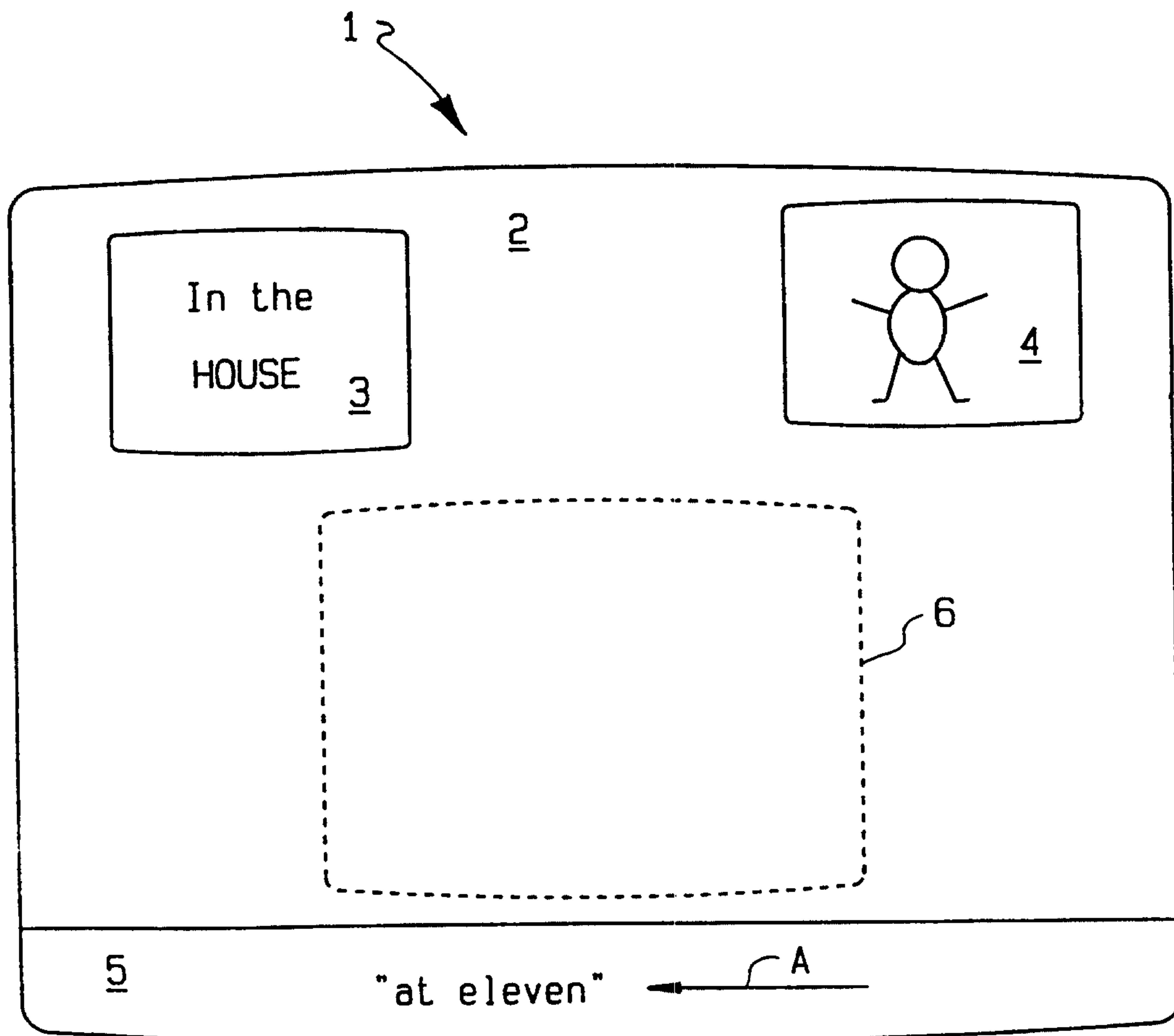


FIG. 1
Prior Art

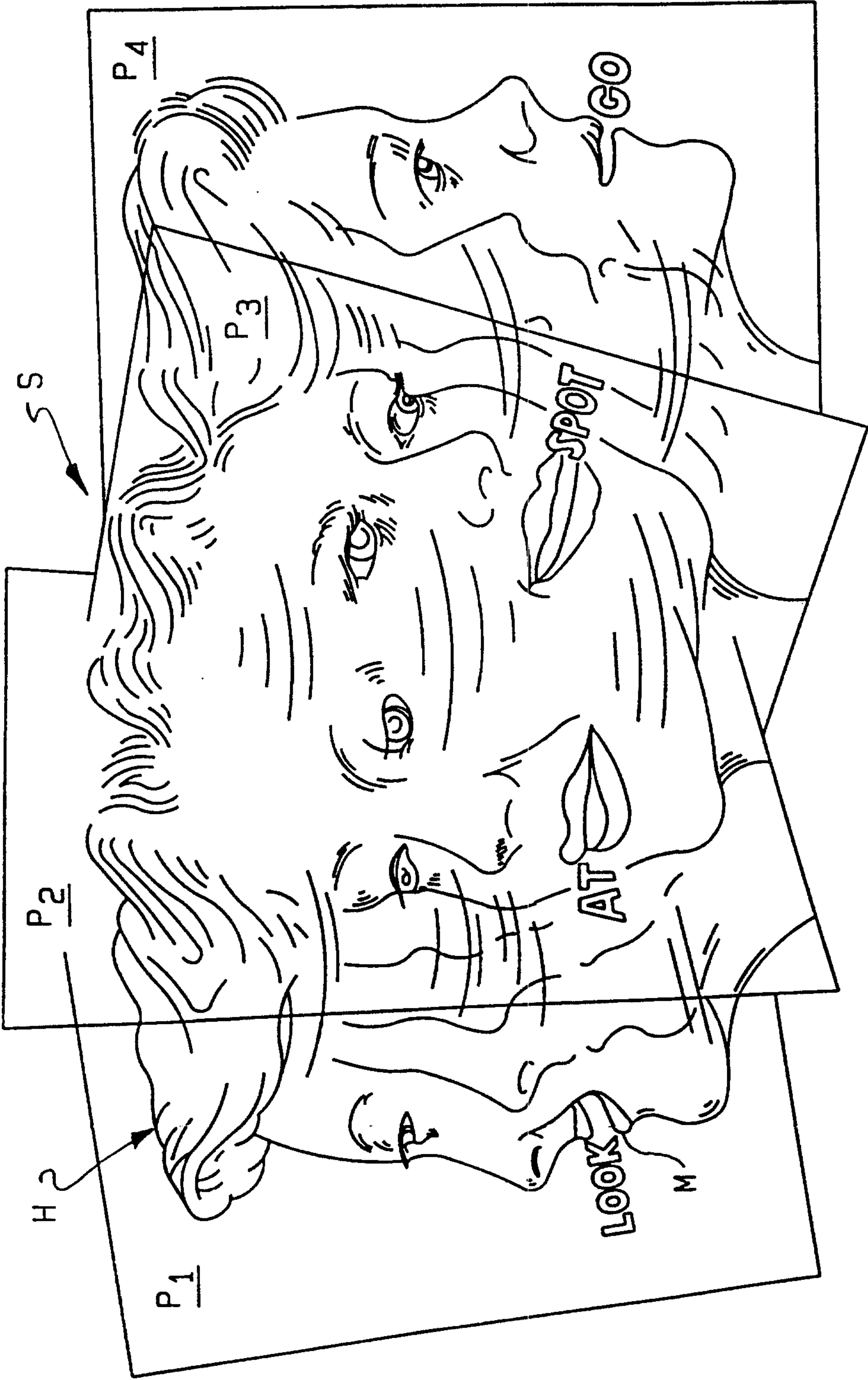


FIG. 2

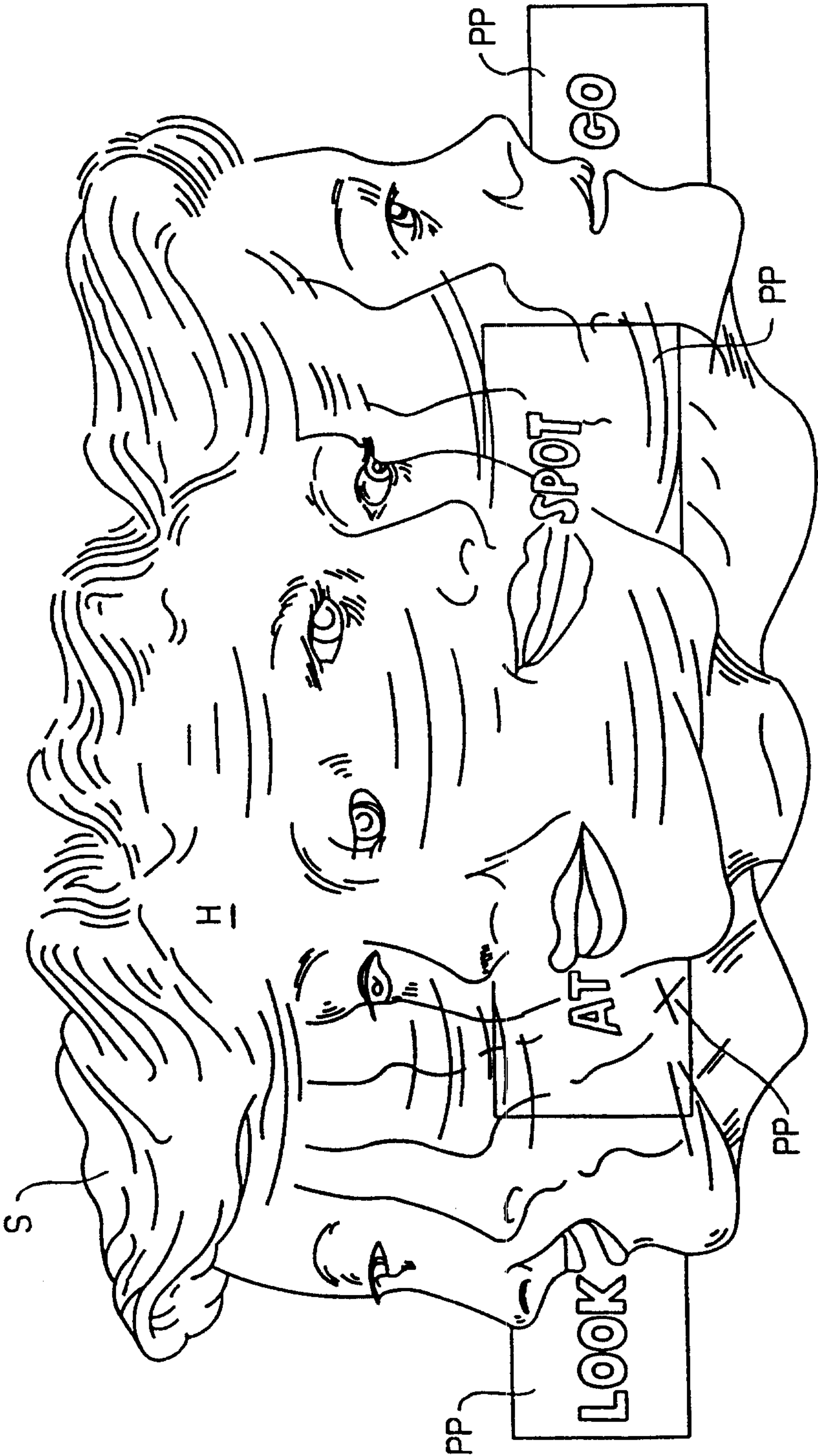
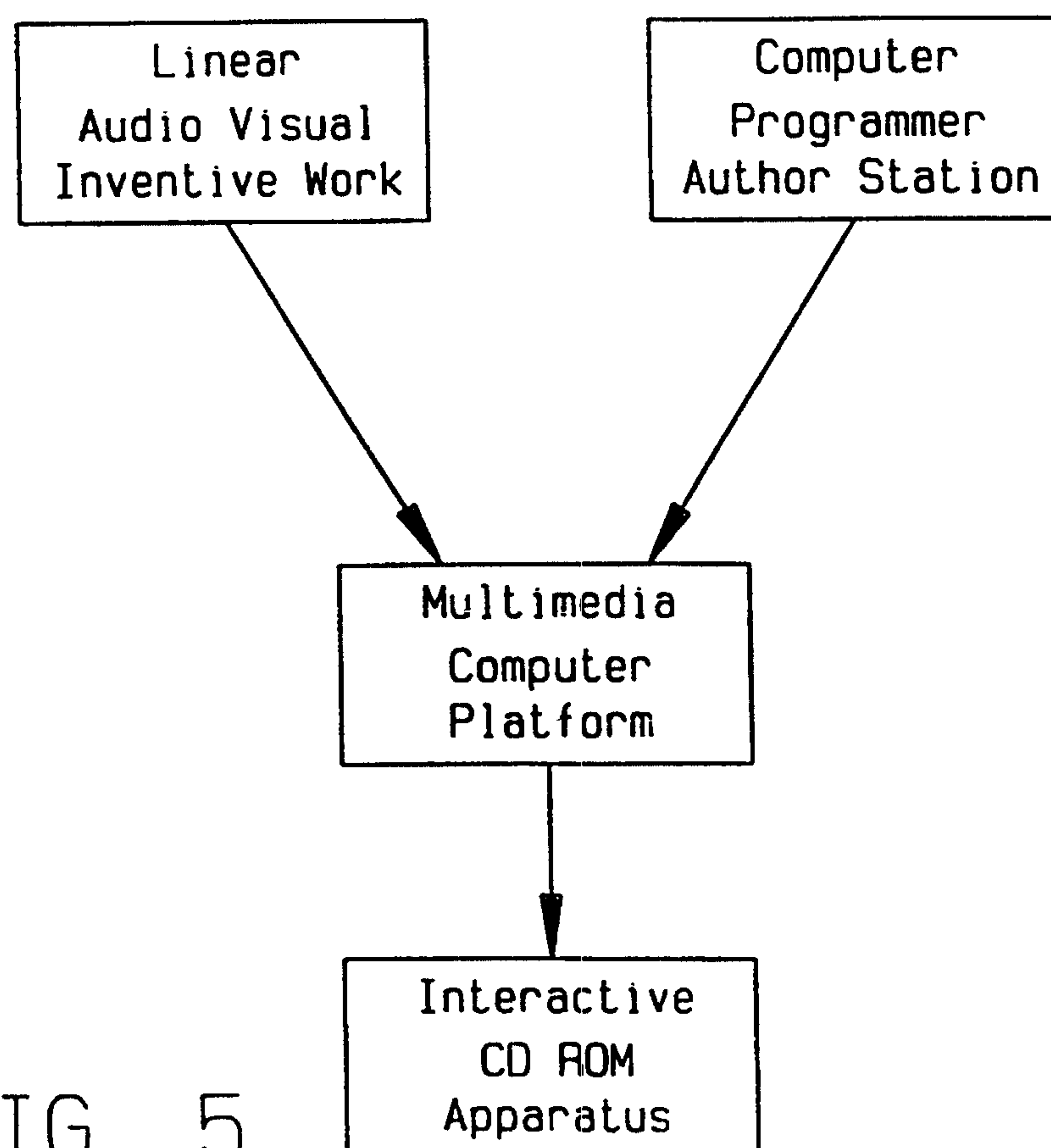
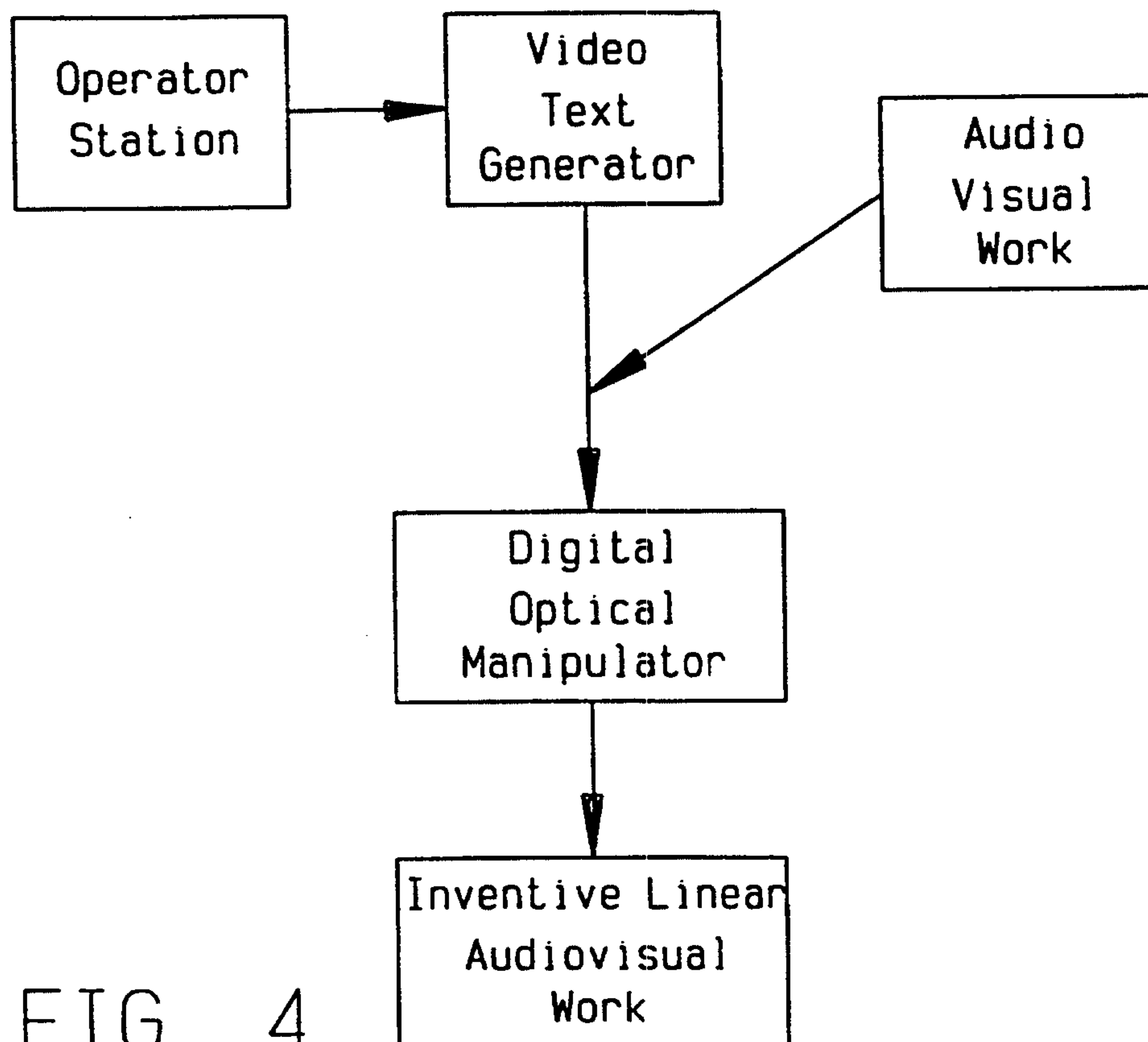


FIG. 3



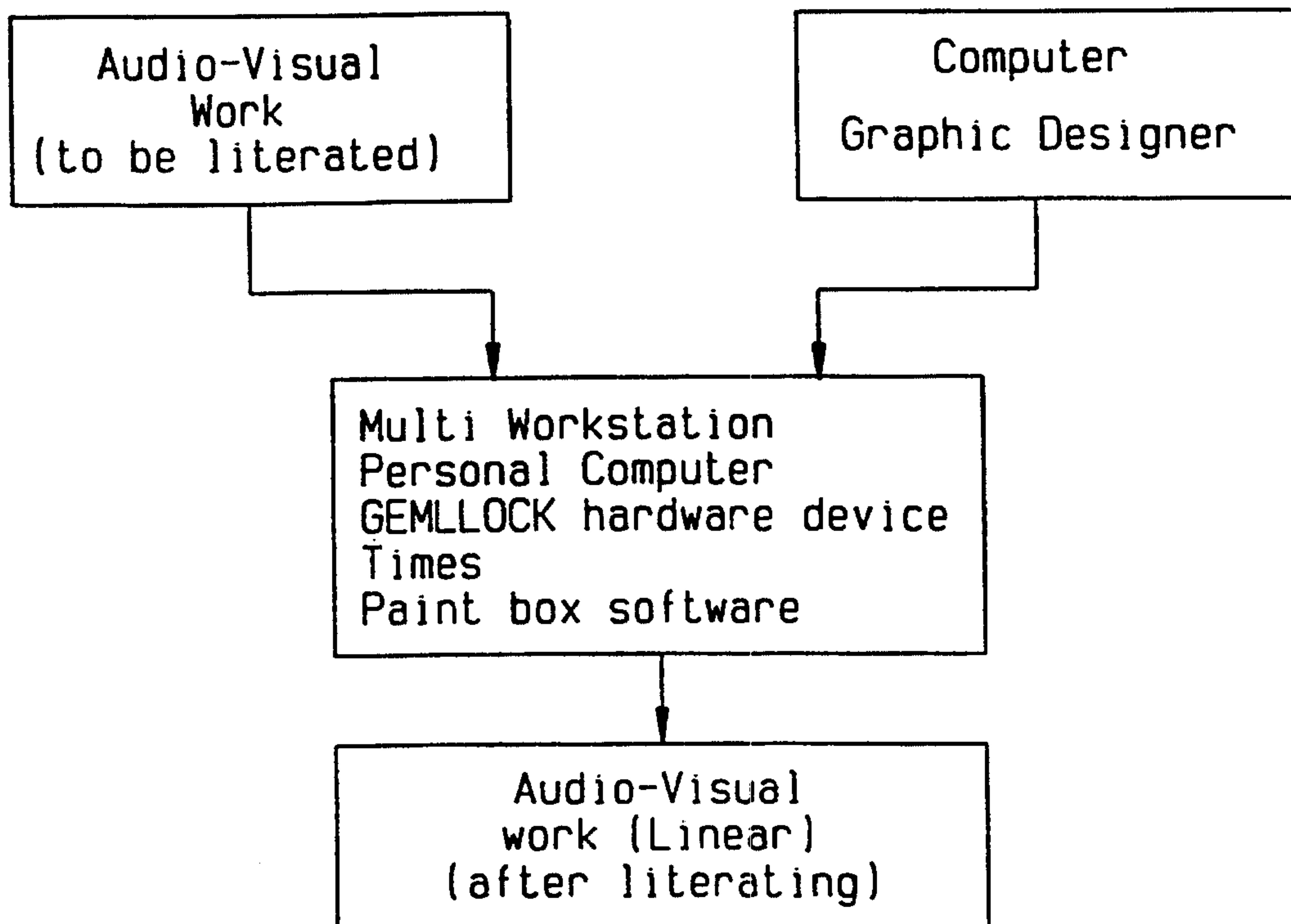


FIG. 6

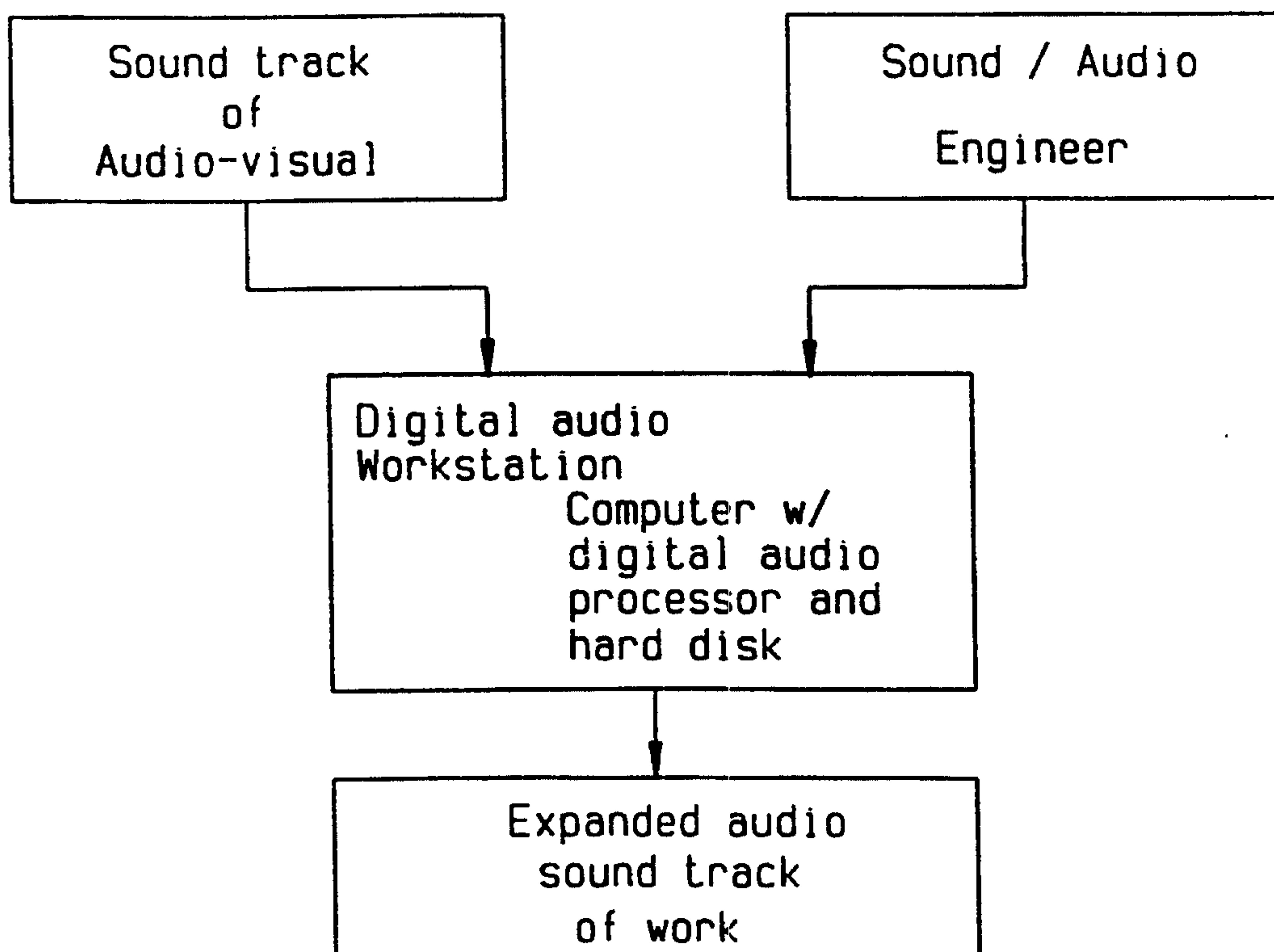


FIG. 7

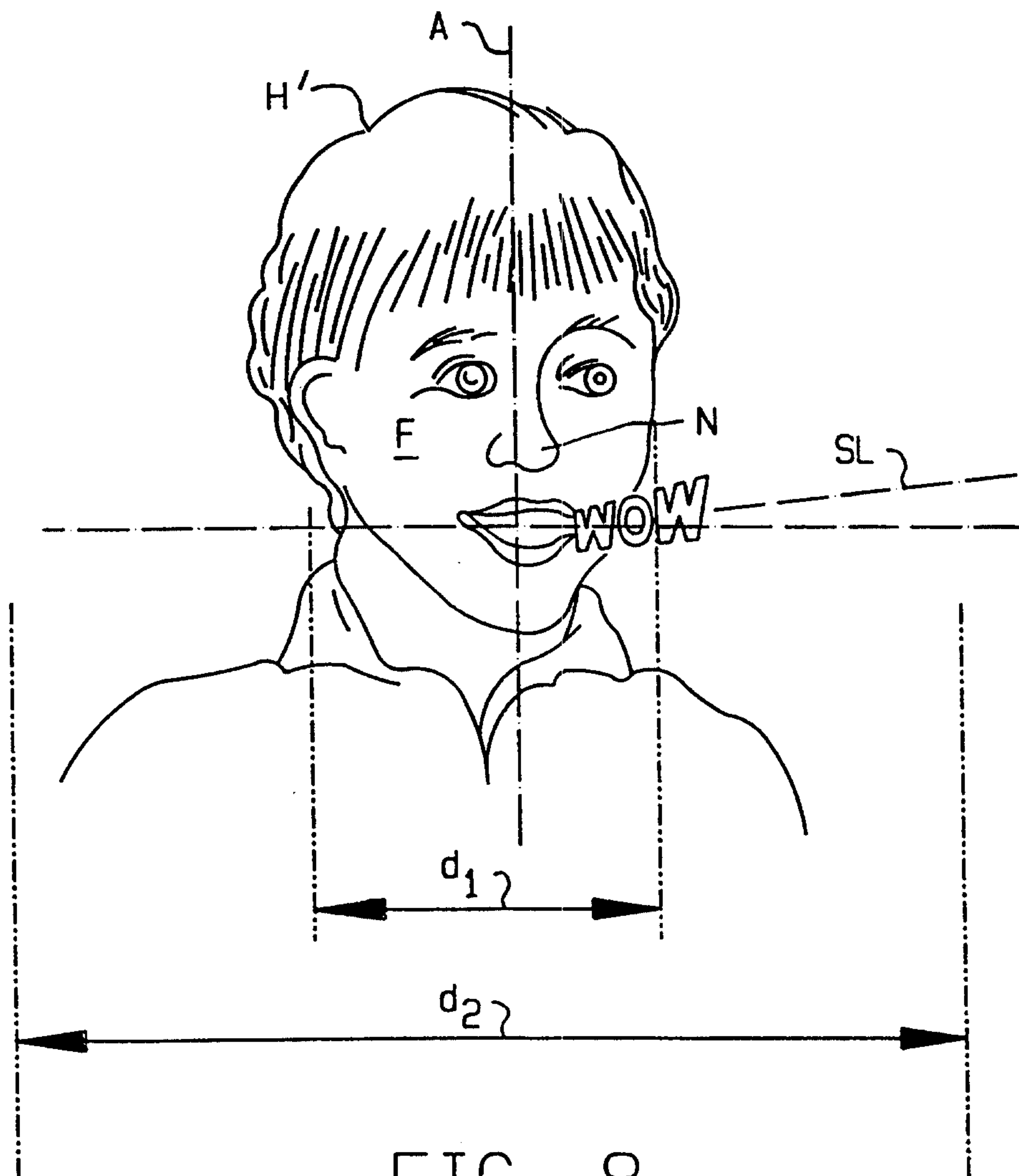


FIG. 8

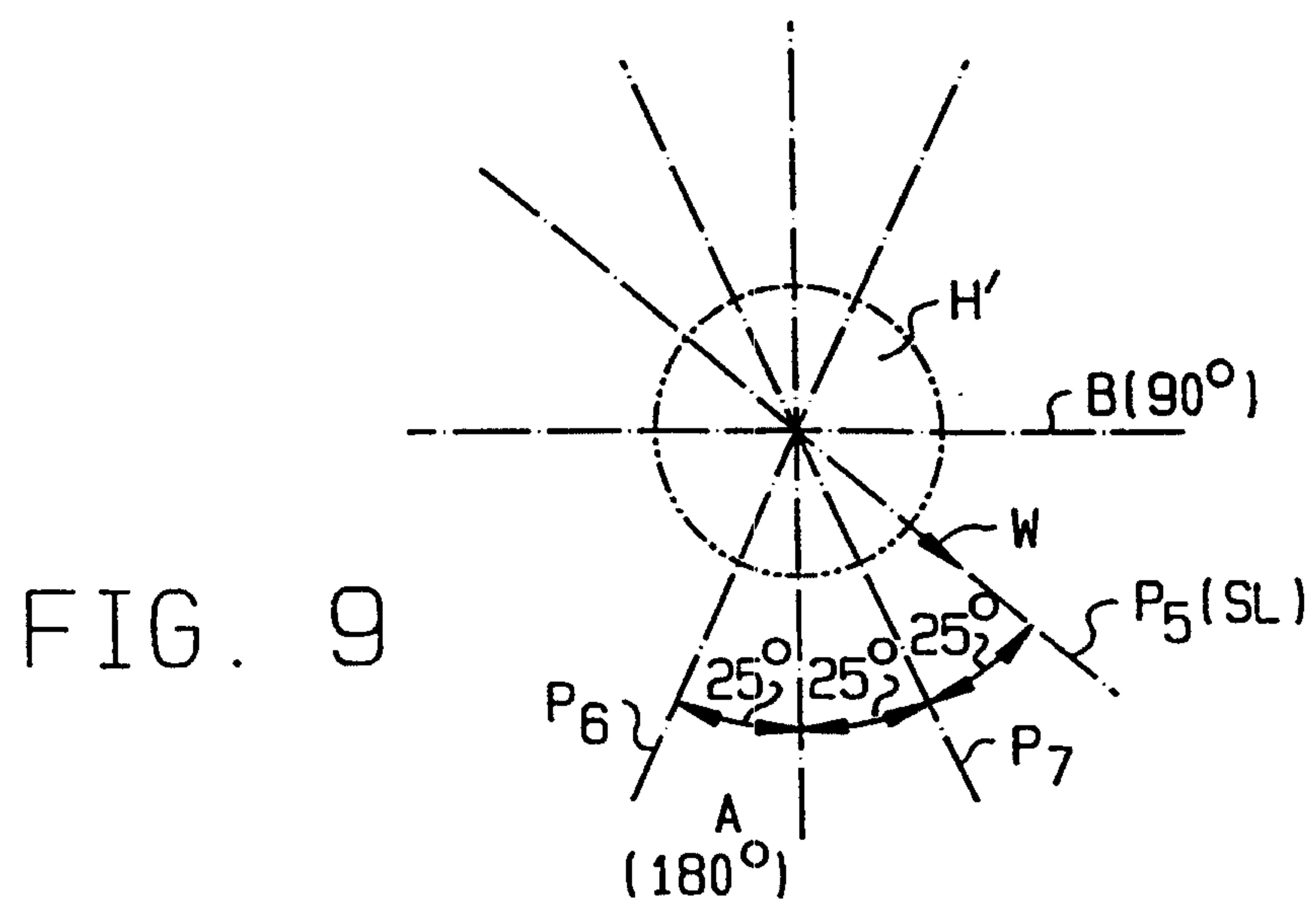


FIG. 9

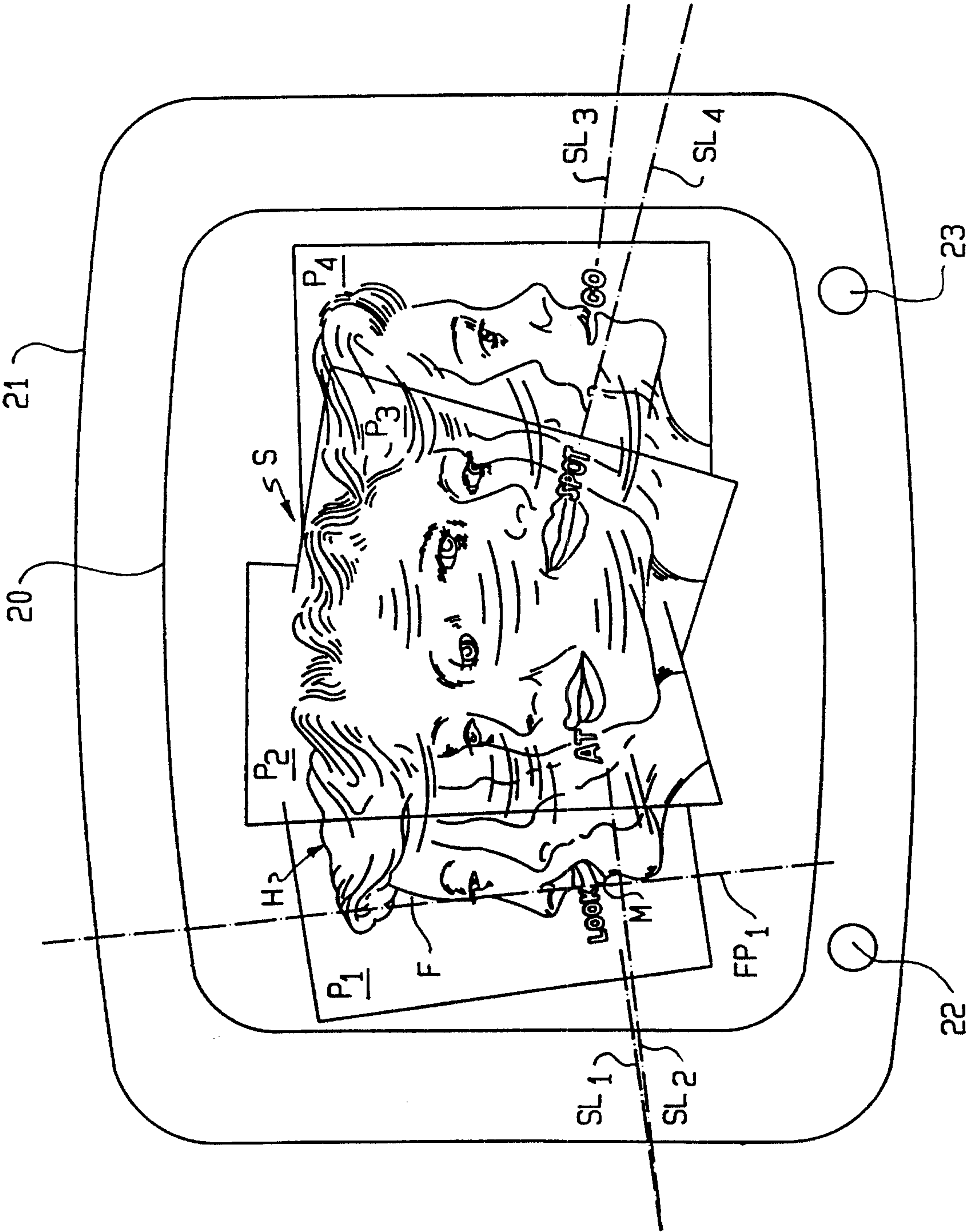


FIG. 10

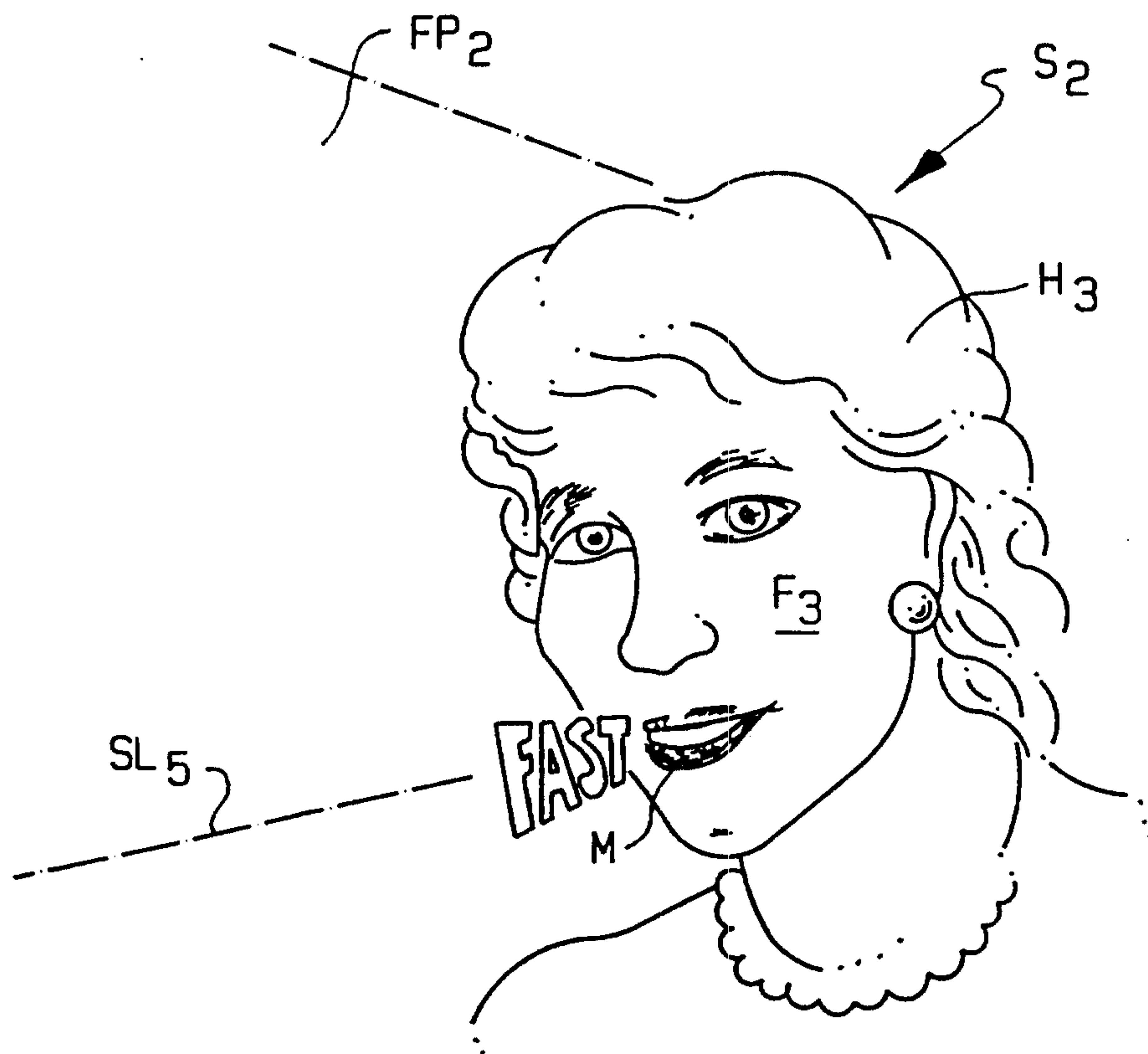


FIG. 11

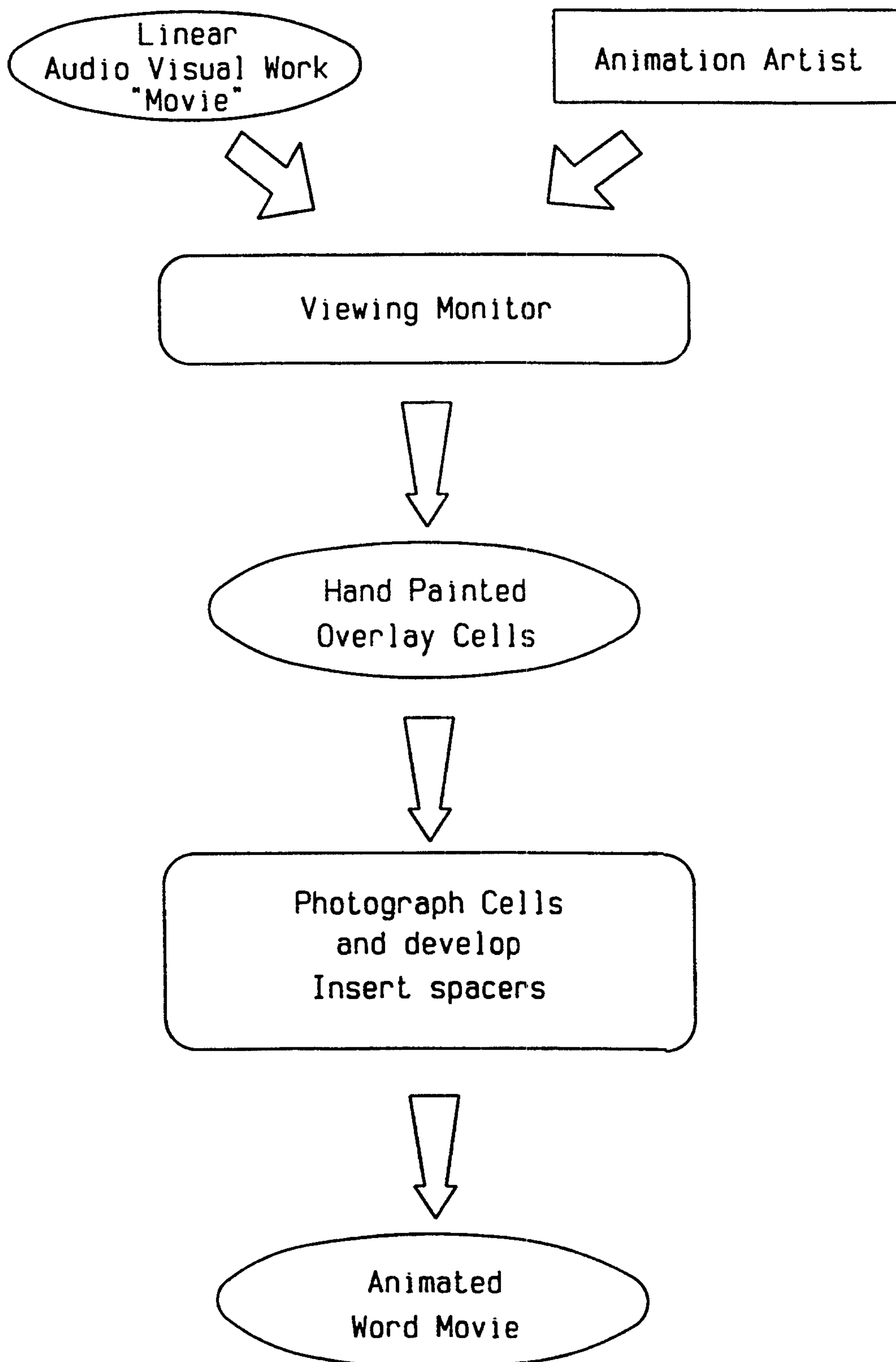


FIG. 12a

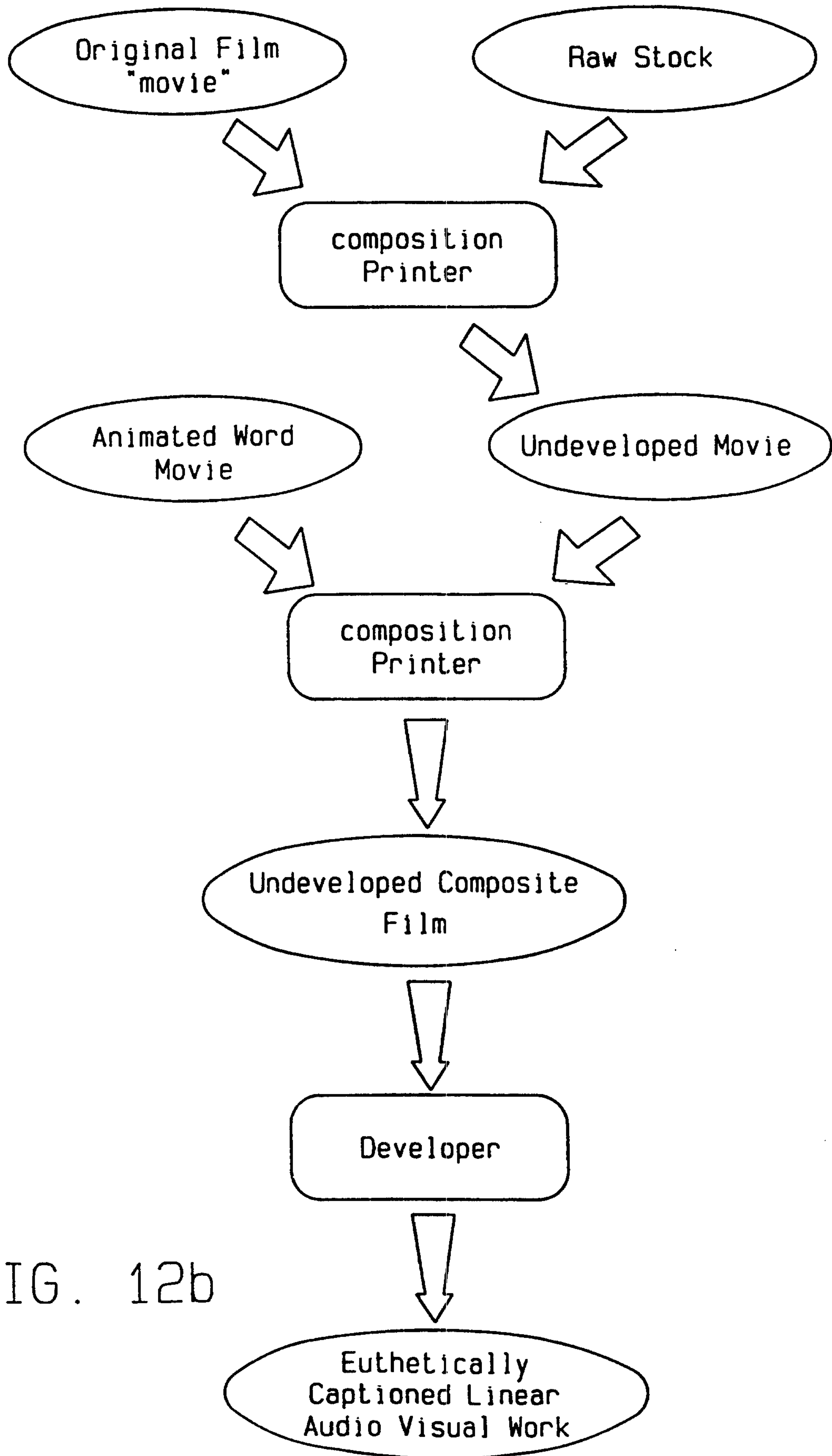


FIG. 12b

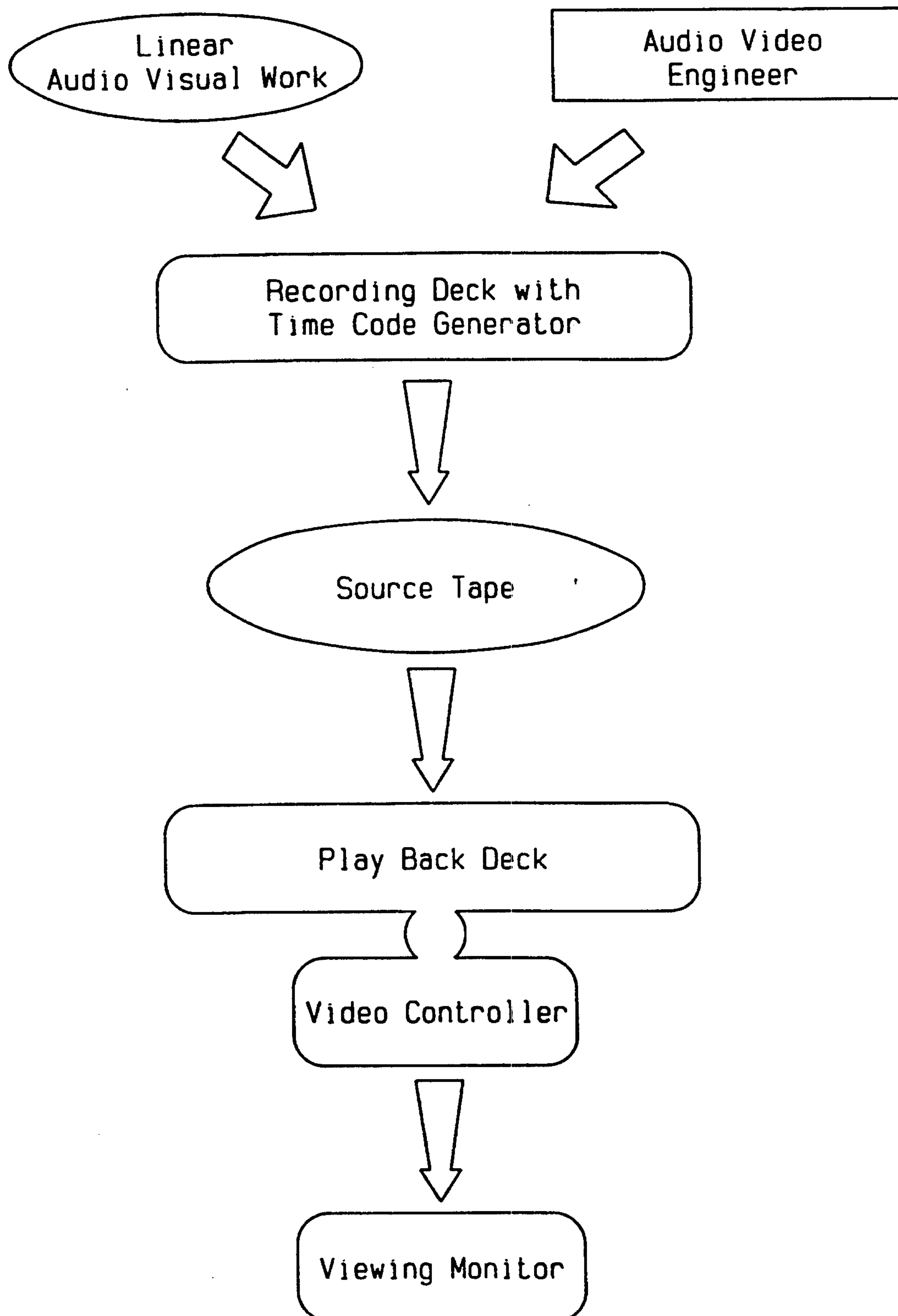


FIG. 13a

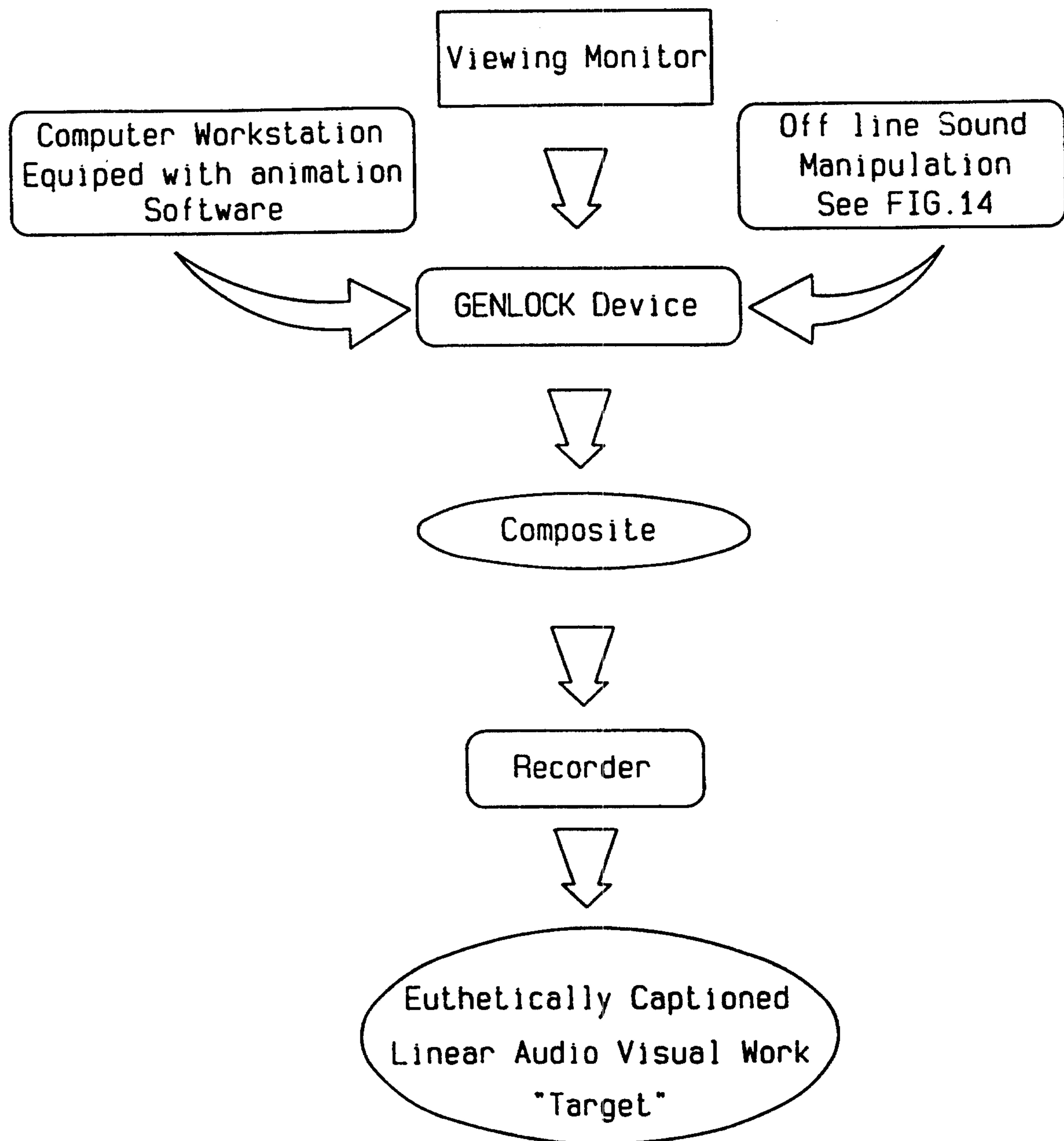


FIG. 13b

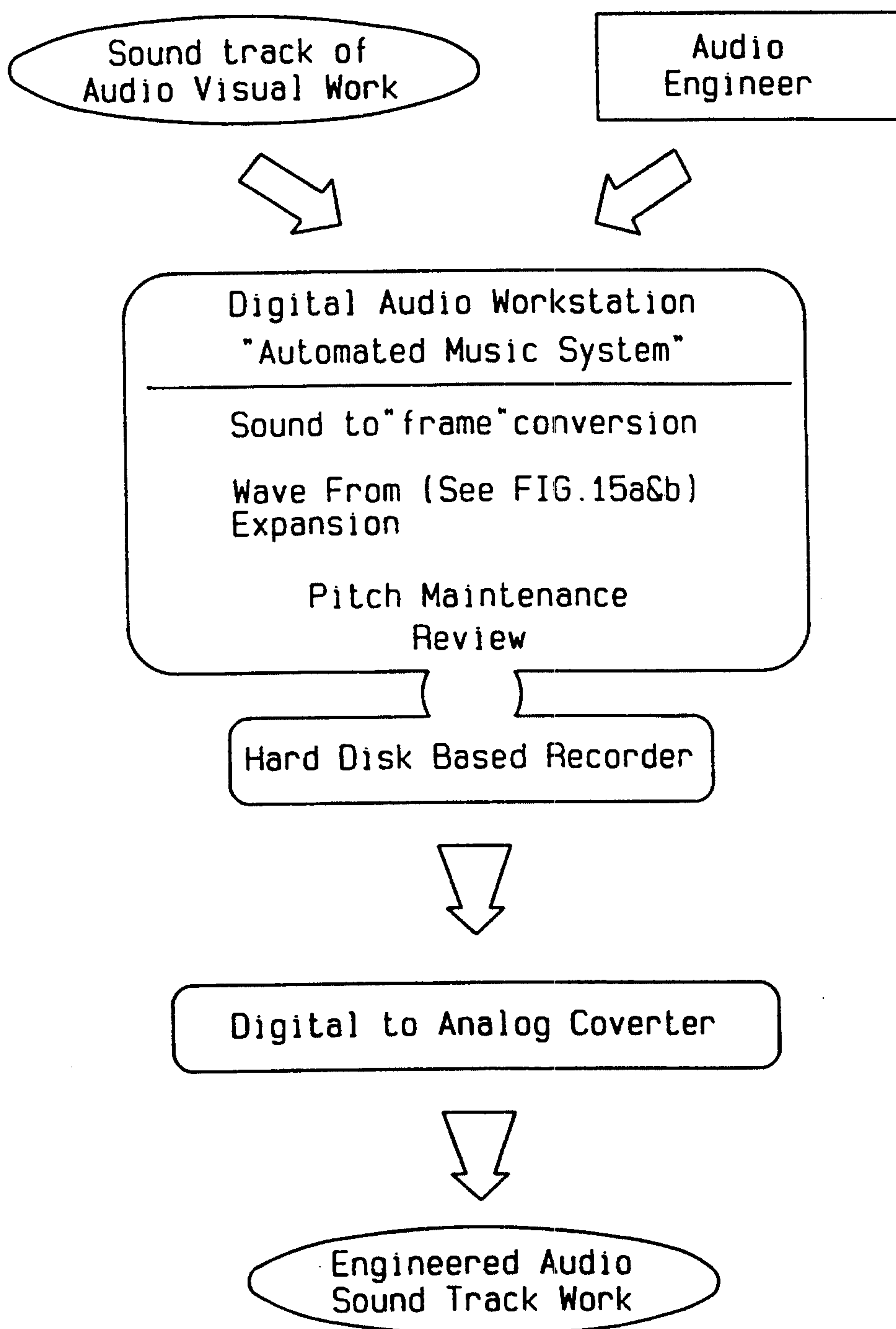


FIG. 14

Original Wave



FIG. 15a

Enhanced Wave

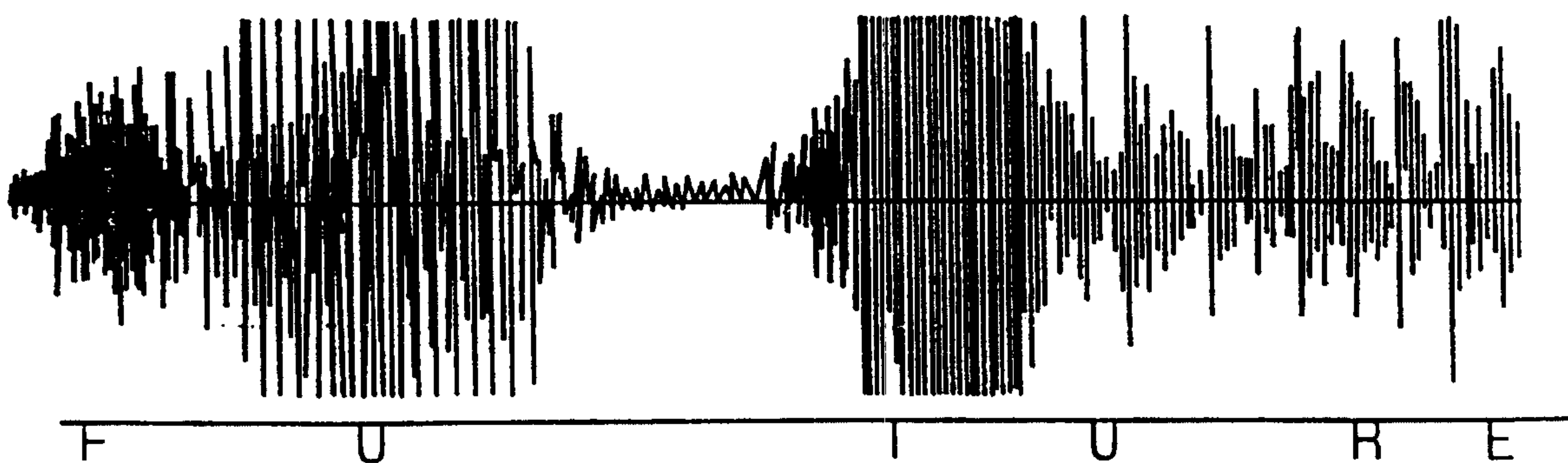


FIG. 15a b

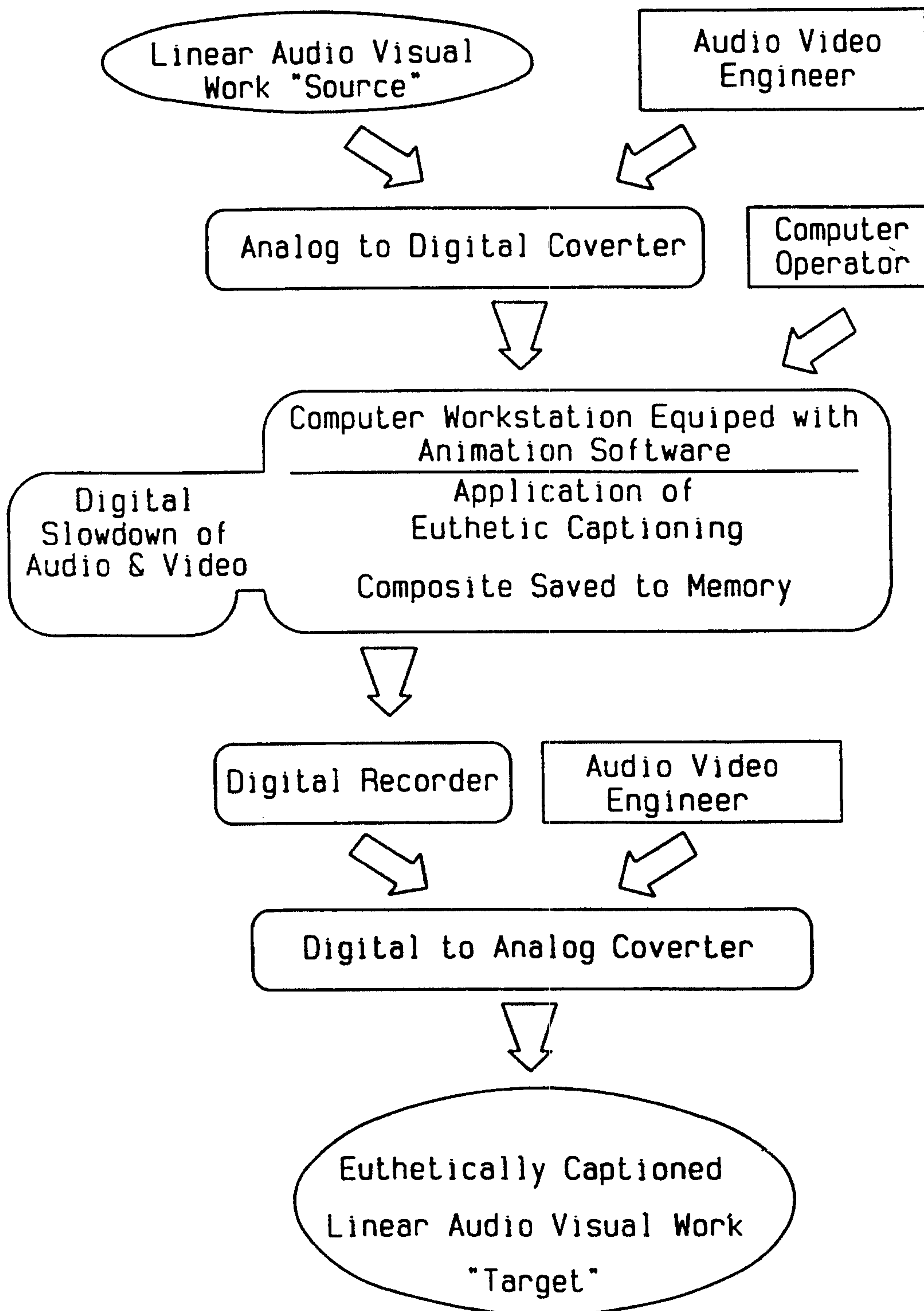


FIG. 16

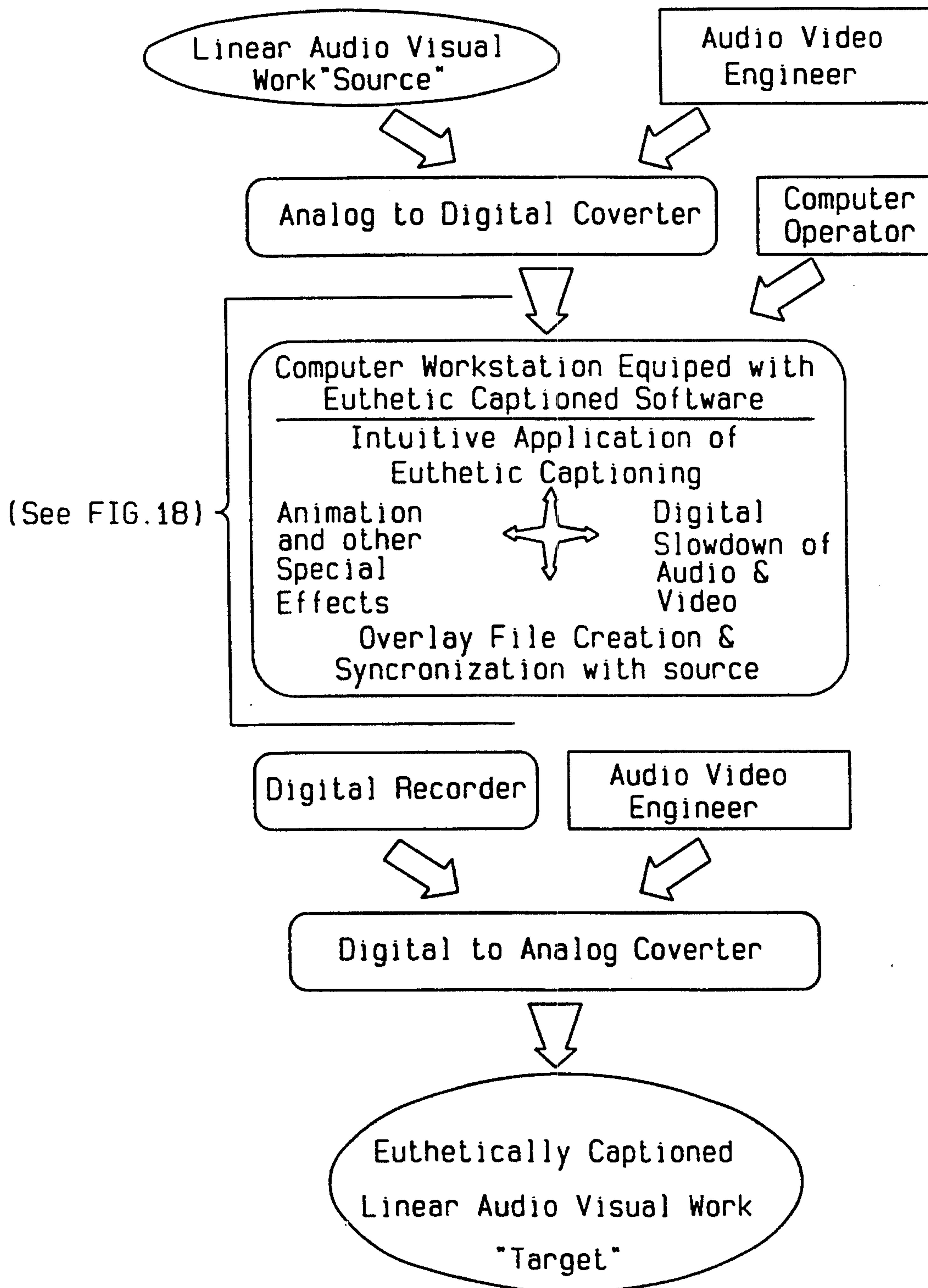


FIG. 17

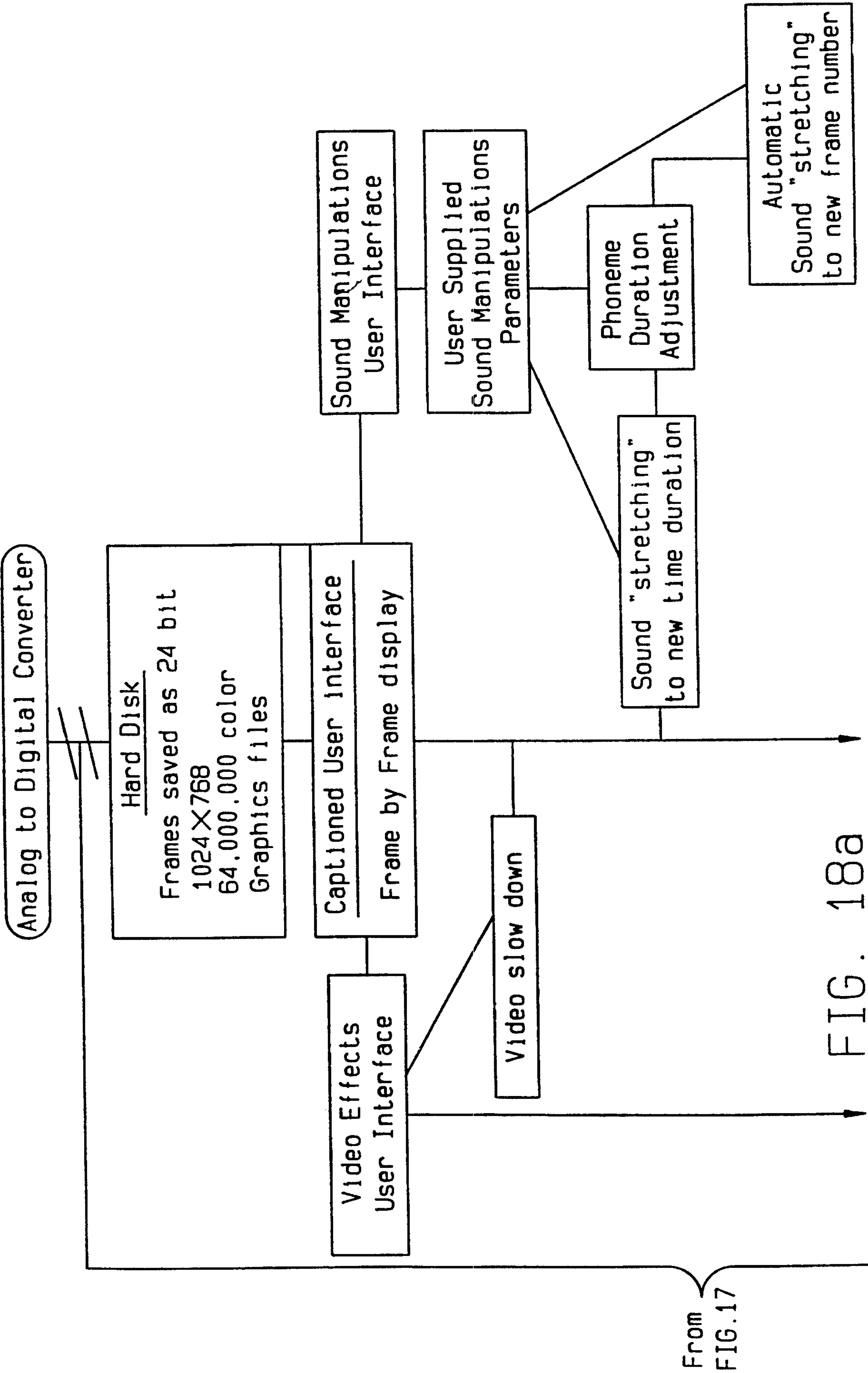


FIG. 18a

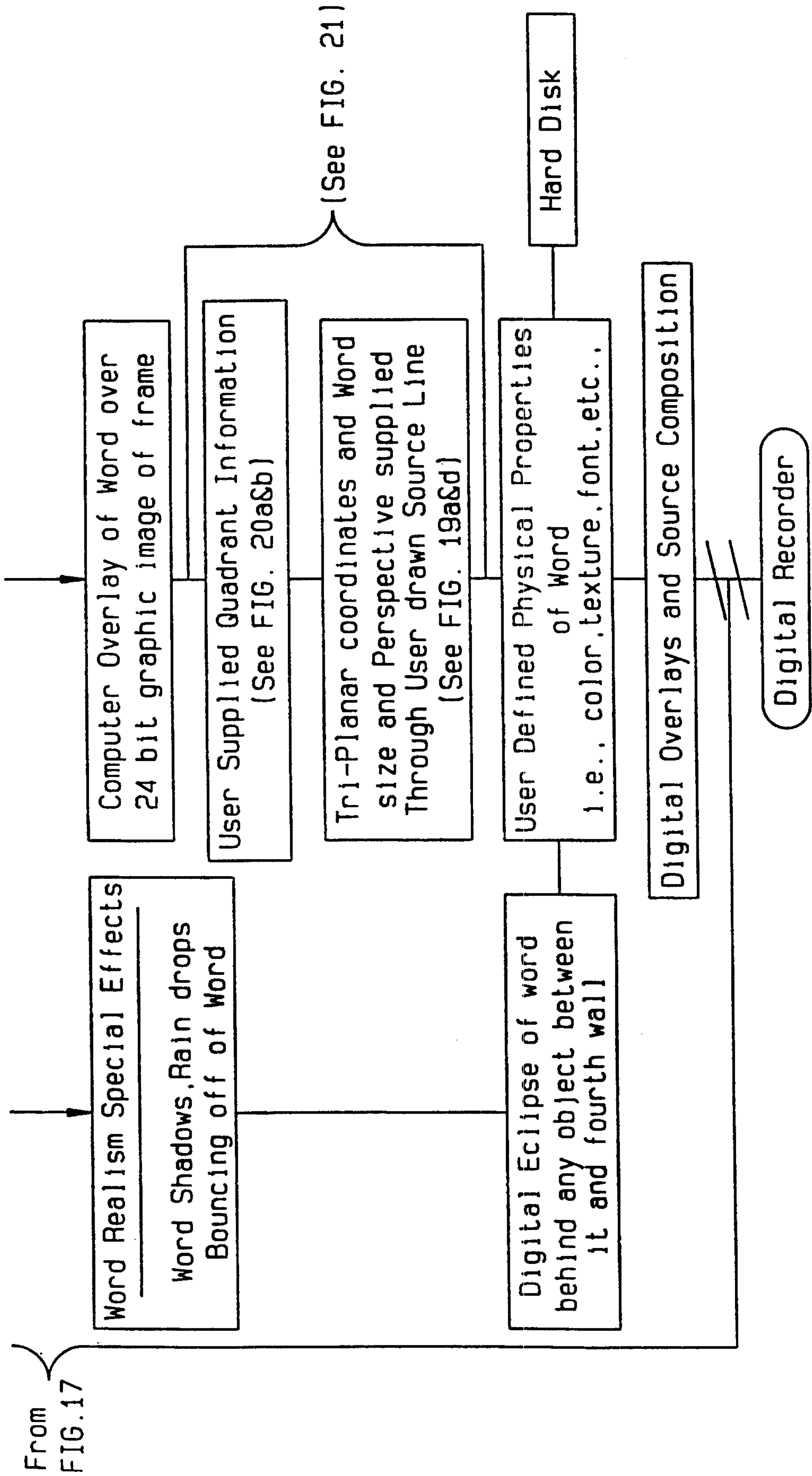


FIG. 18b

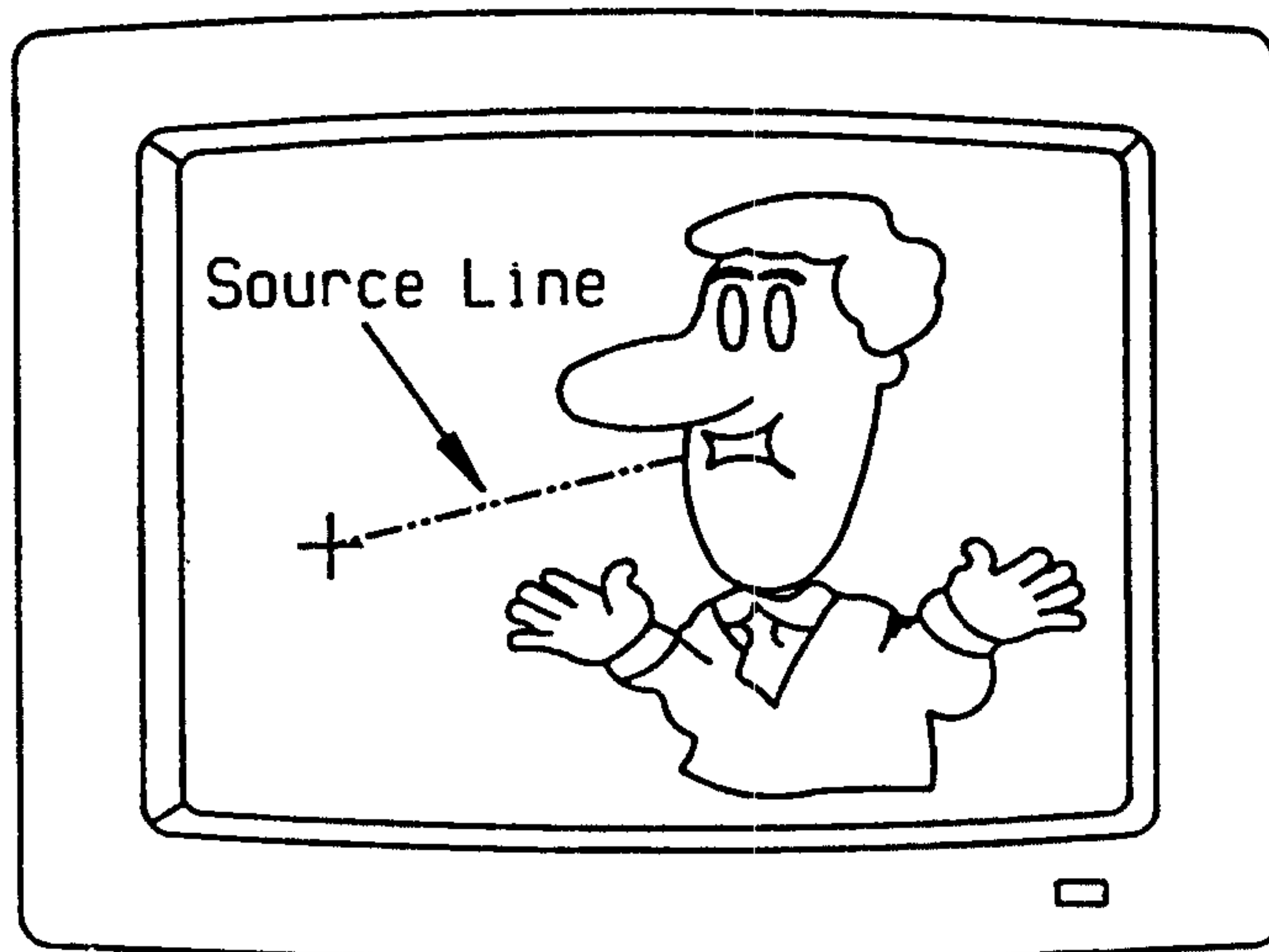


FIG. 19a

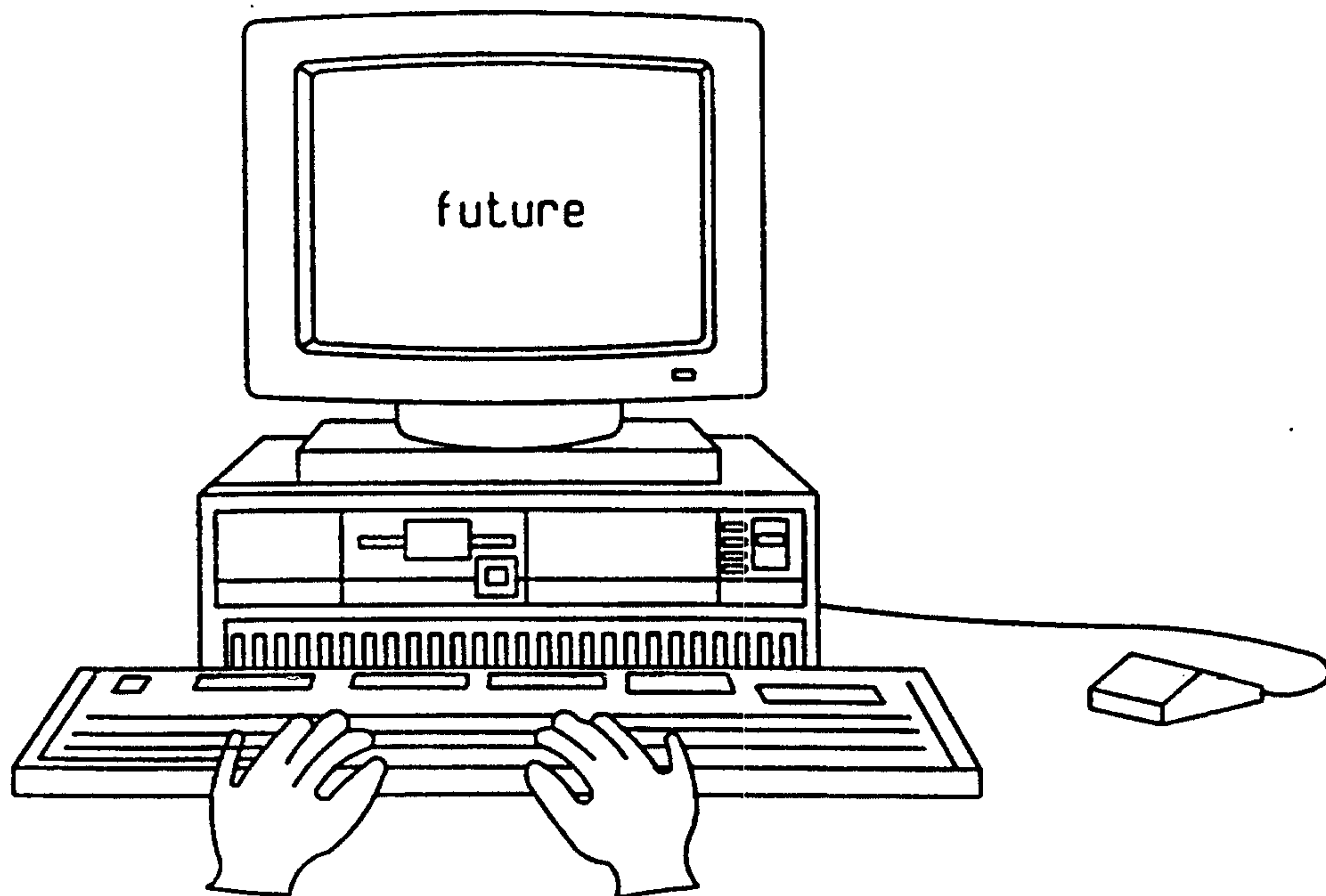


FIG. 19b

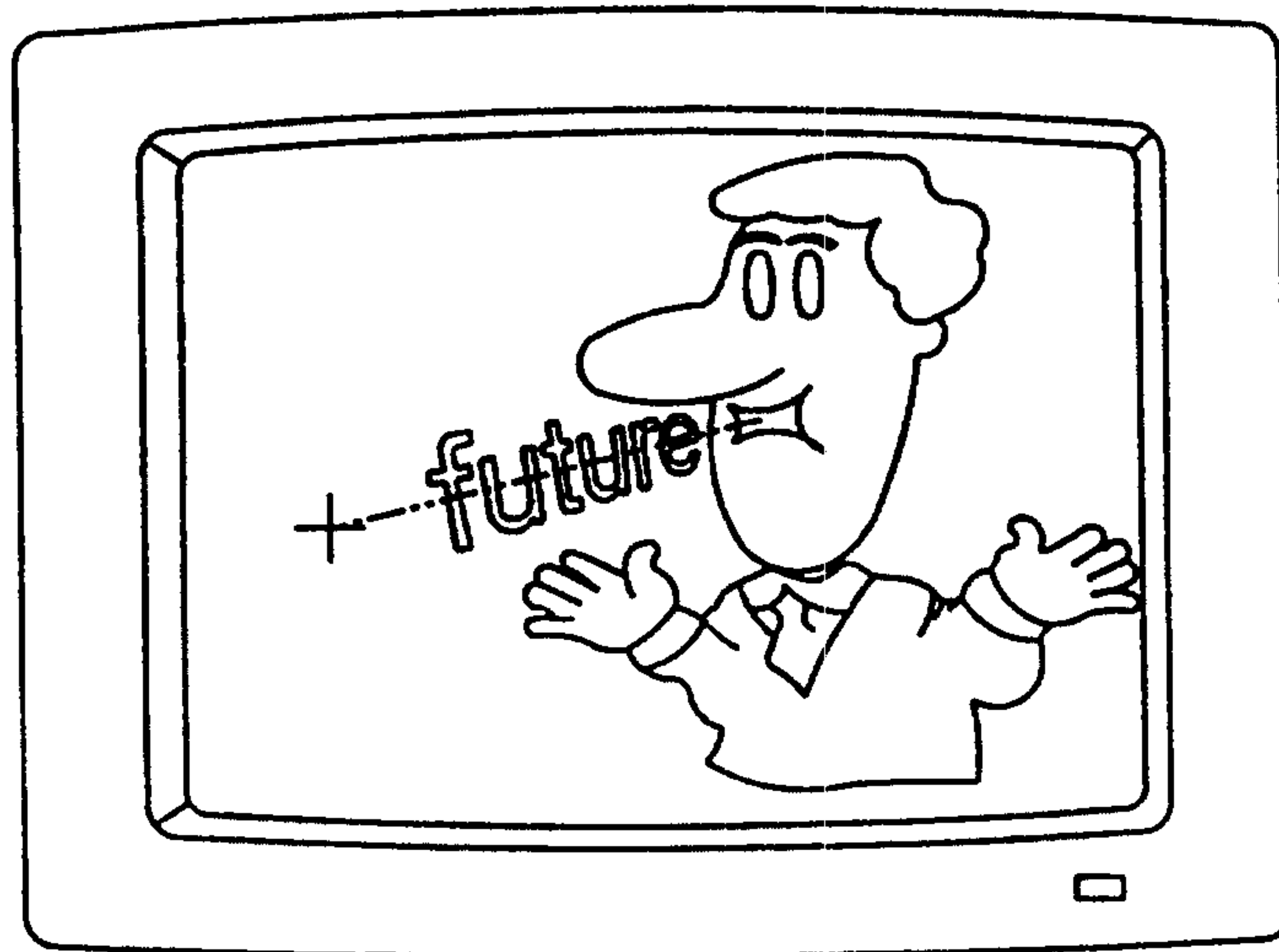


FIG. 19c

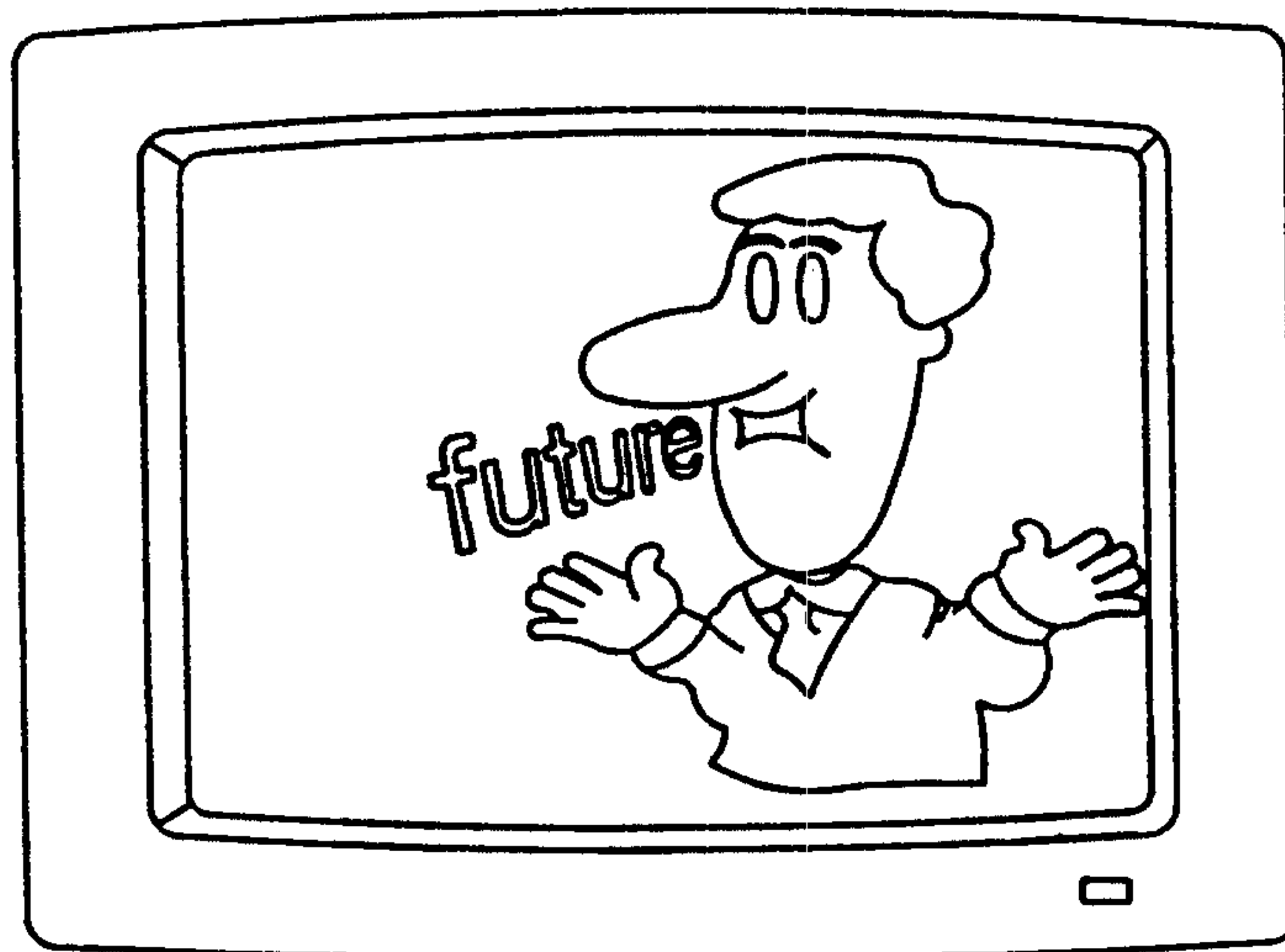


FIG. 19d

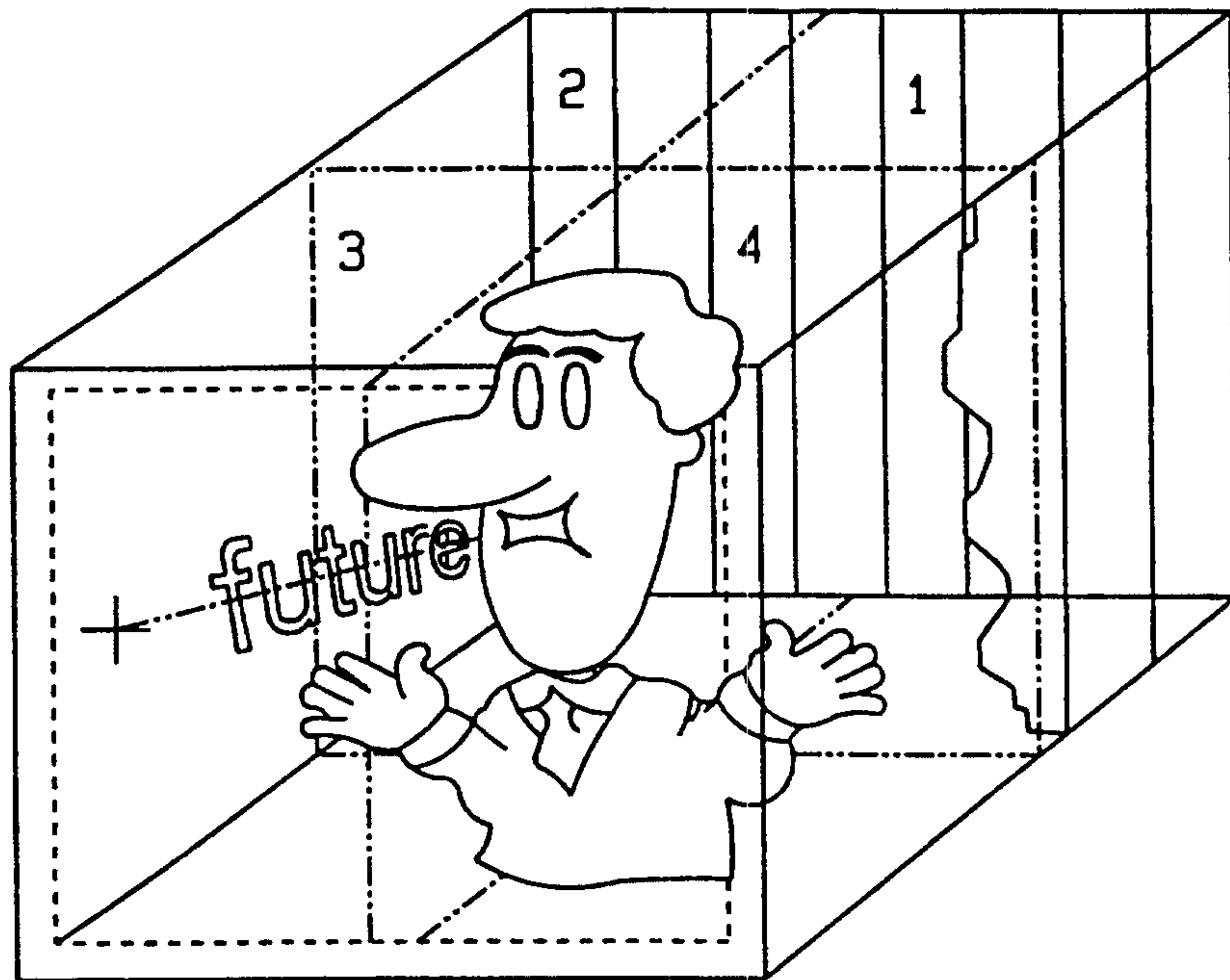


FIG. 20a

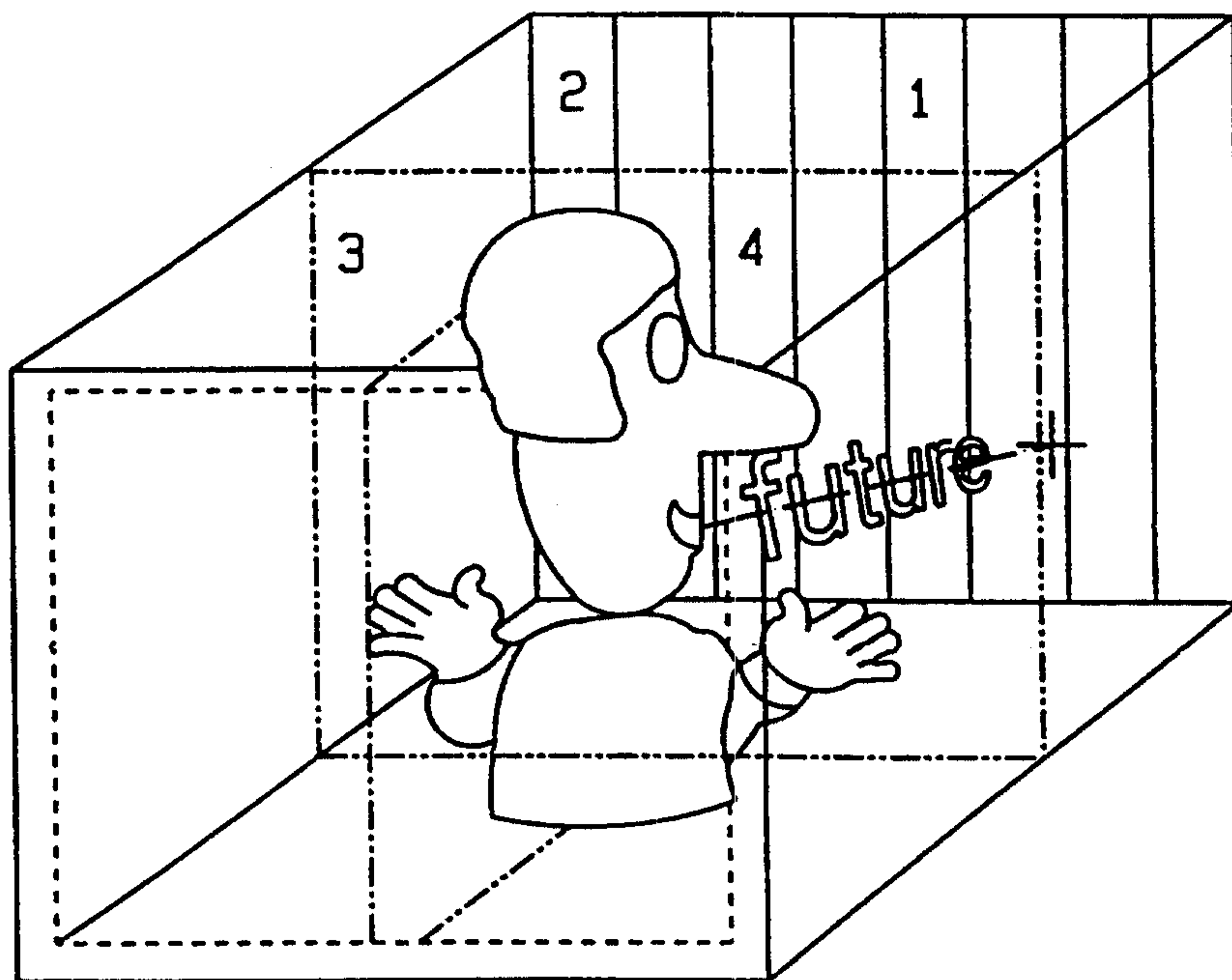


FIG. 20b

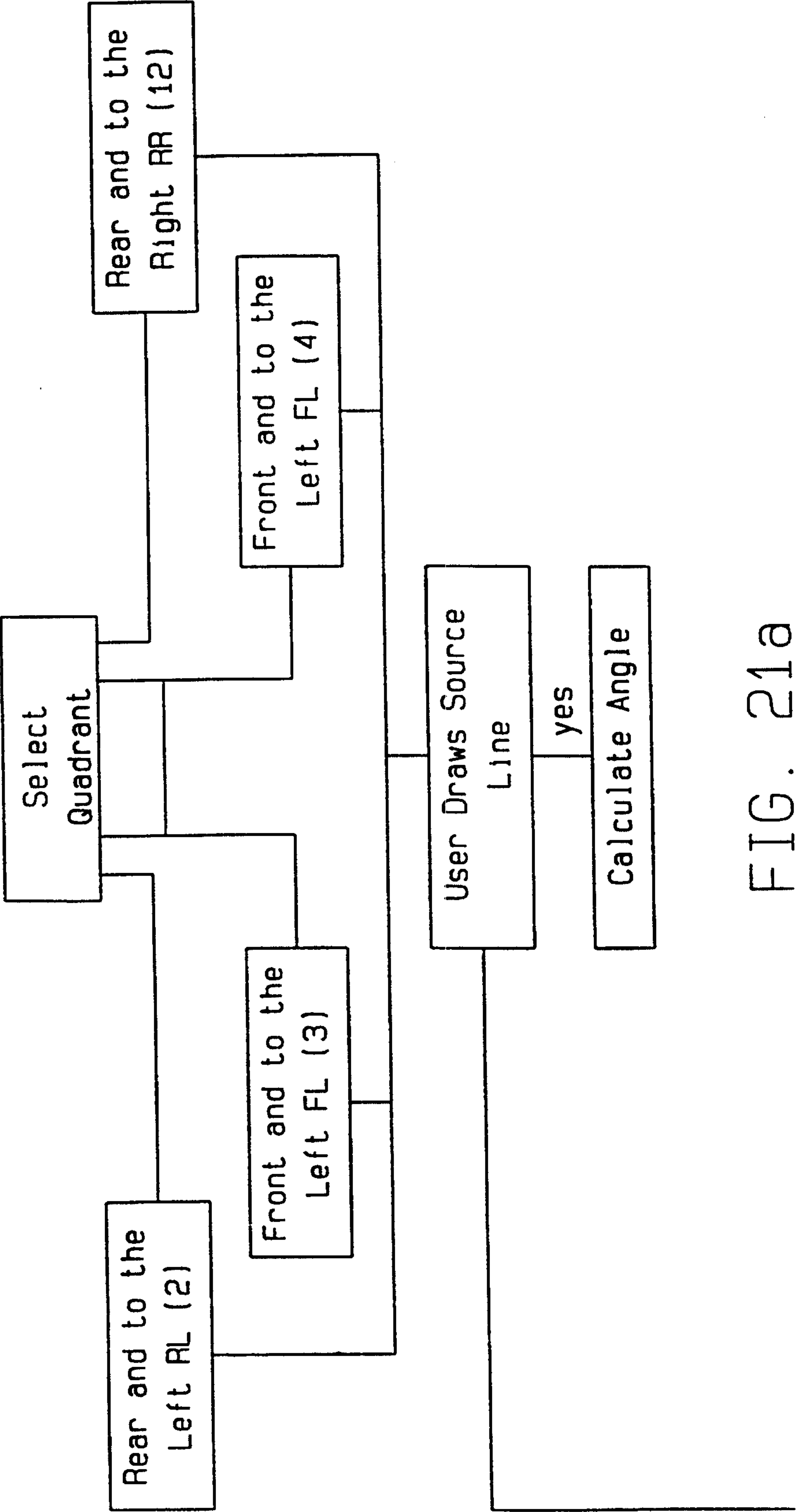


FIG. 21a

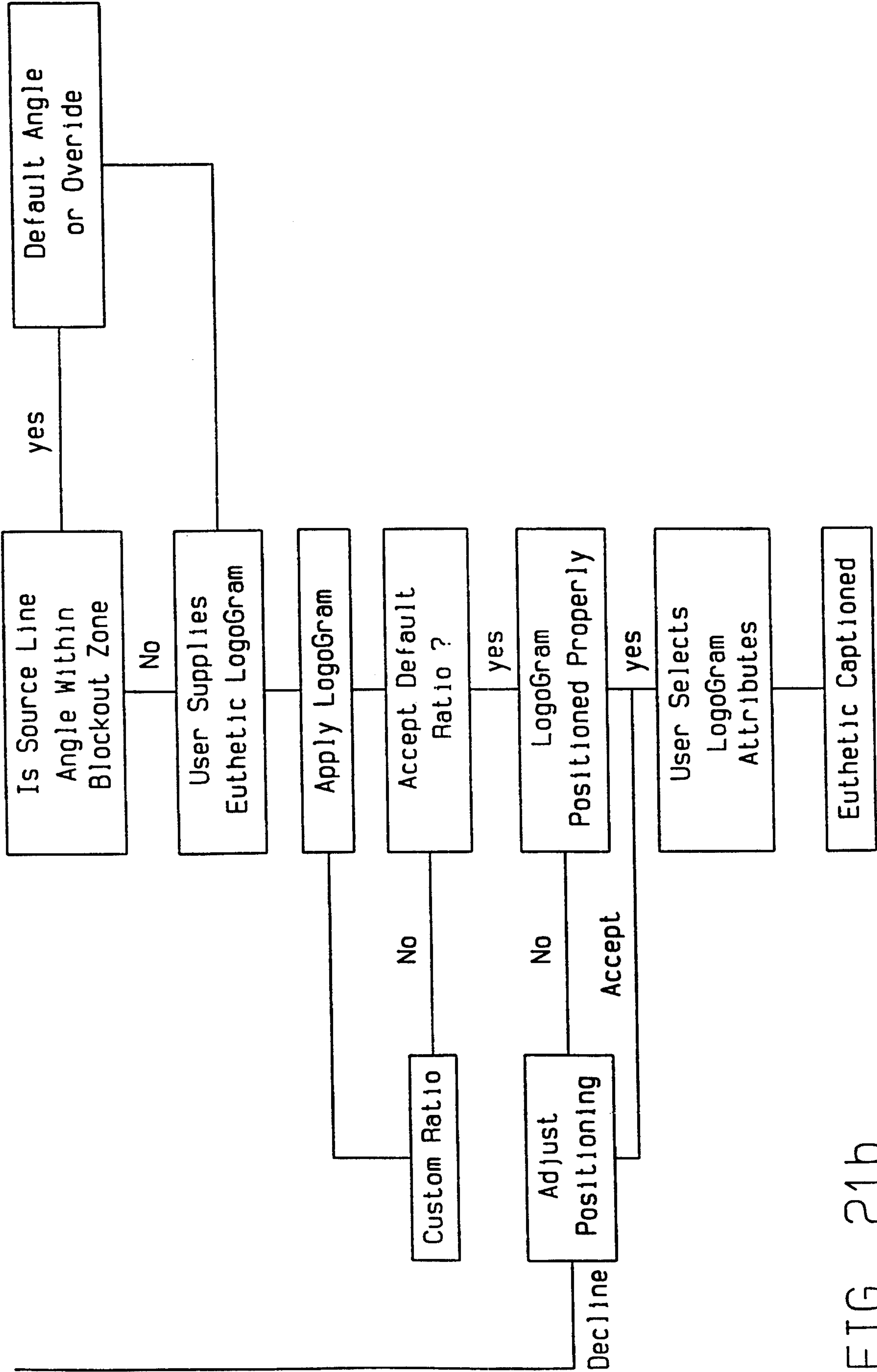


FIG. 21b

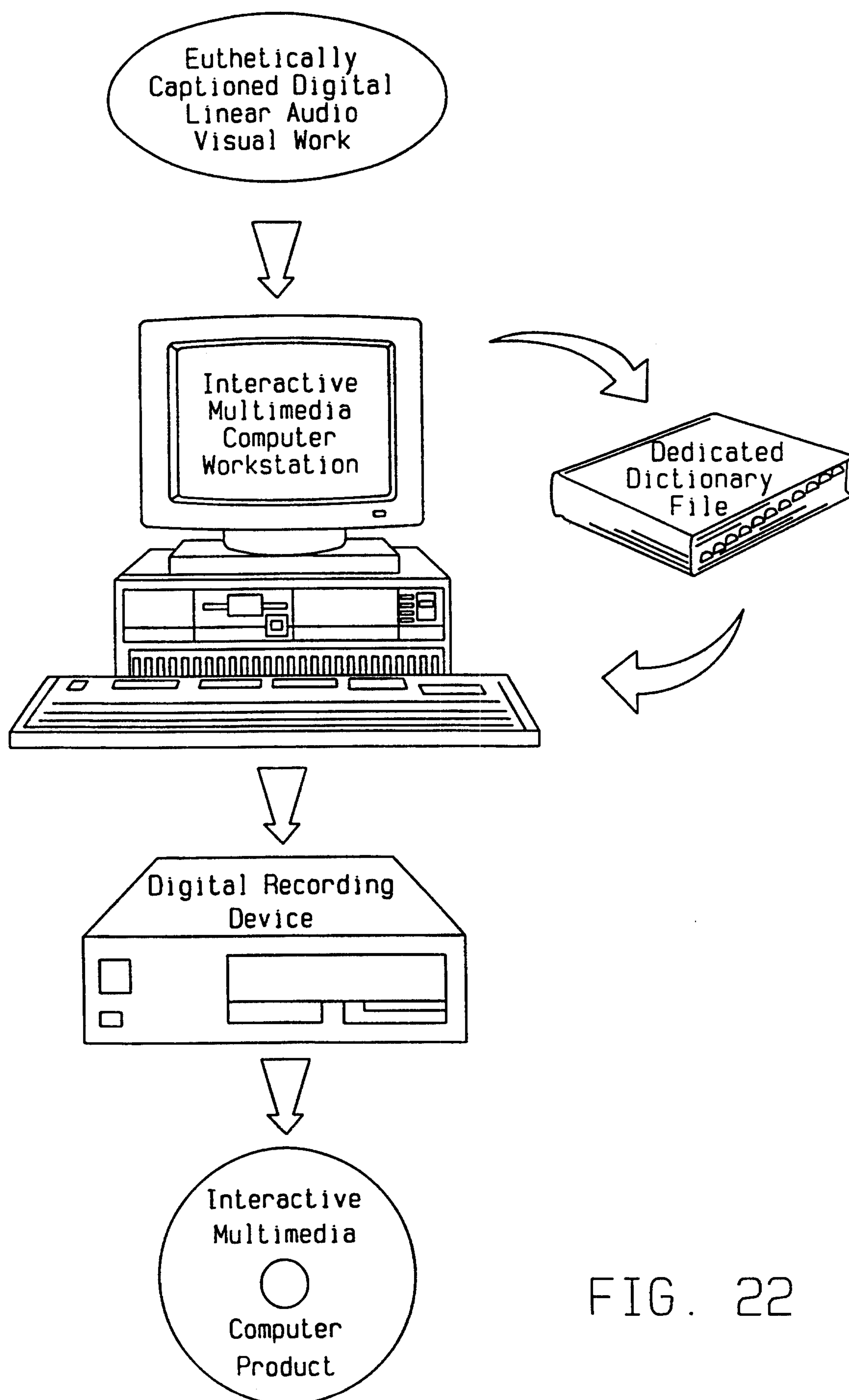


FIG. 22

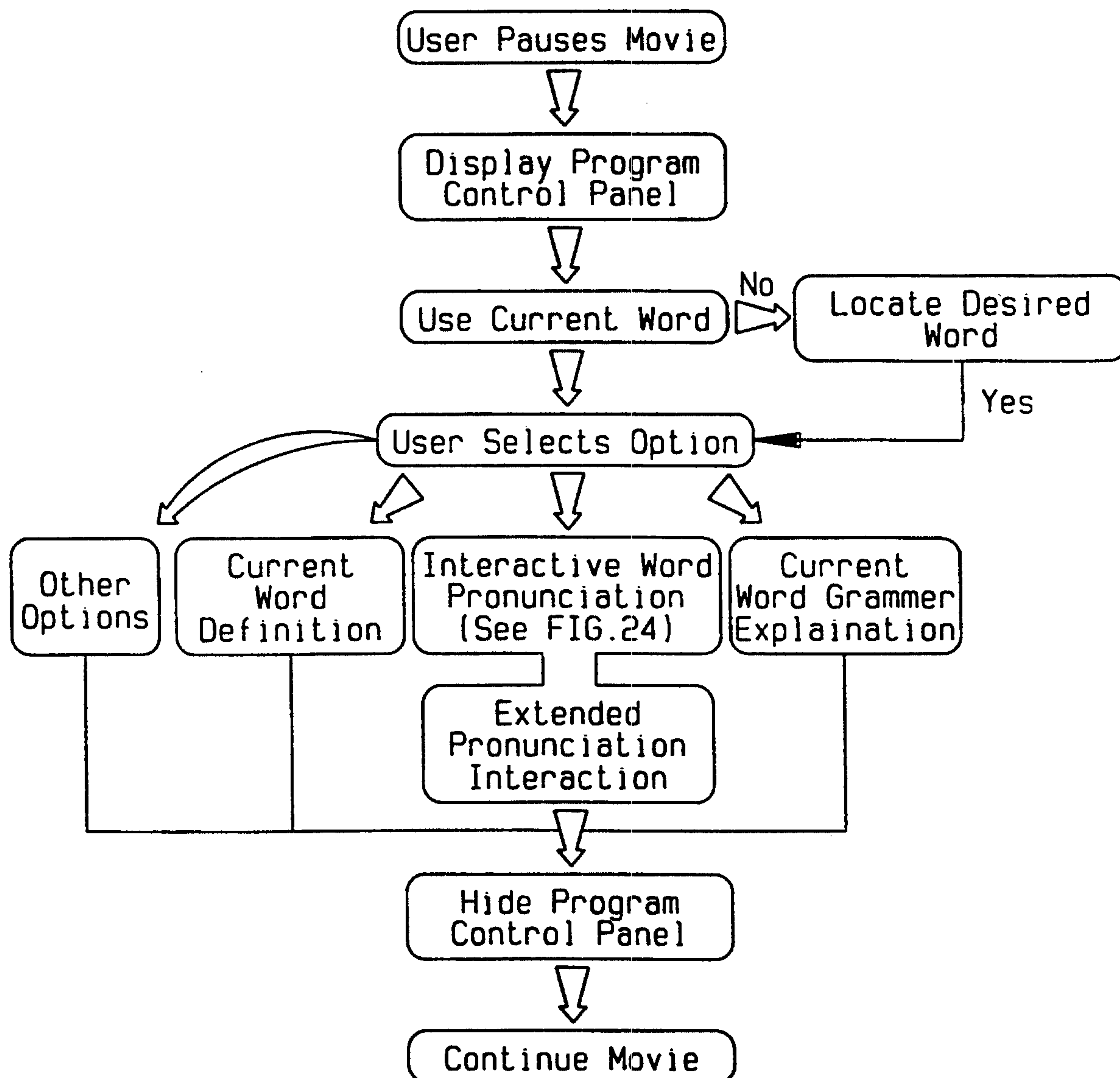


FIG. 23

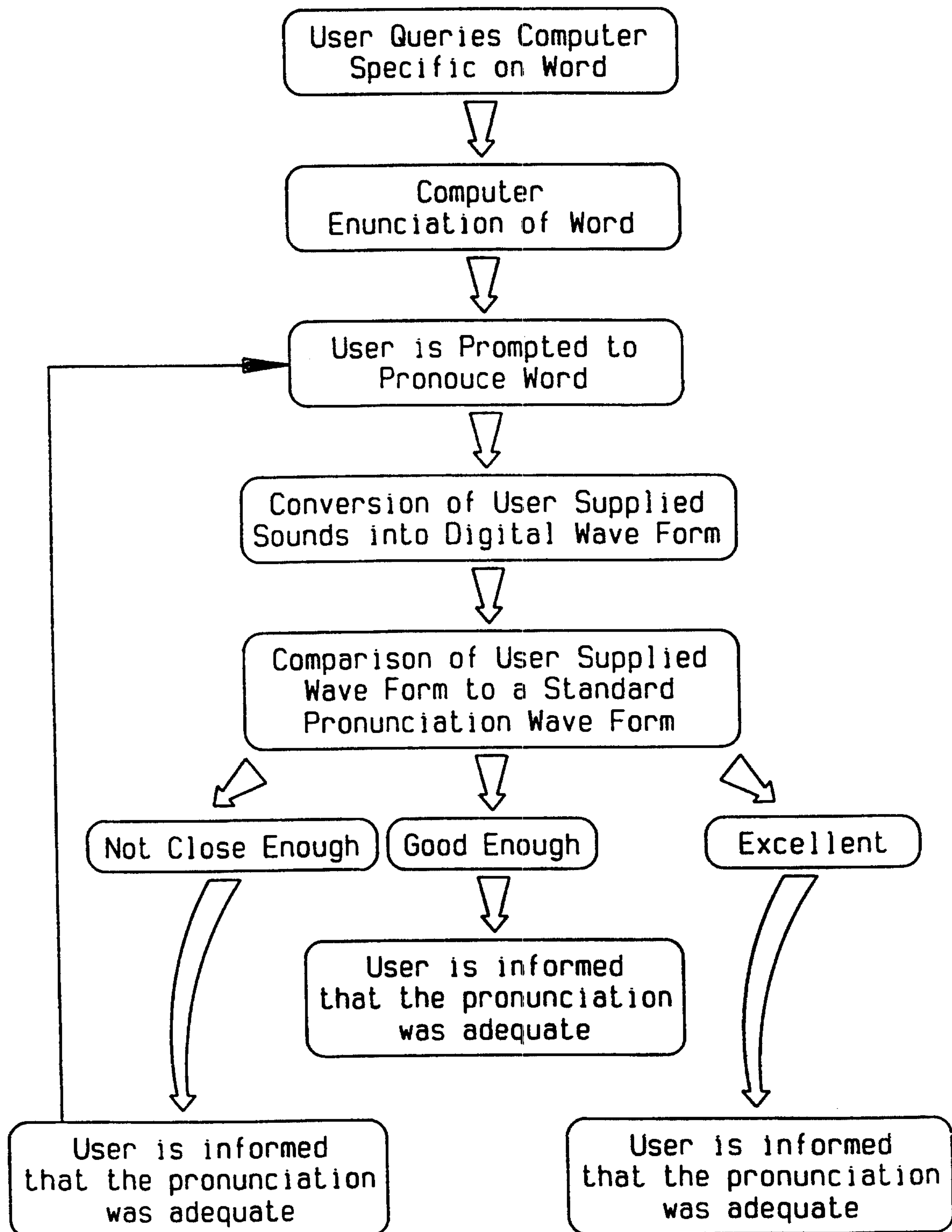


FIG. 24

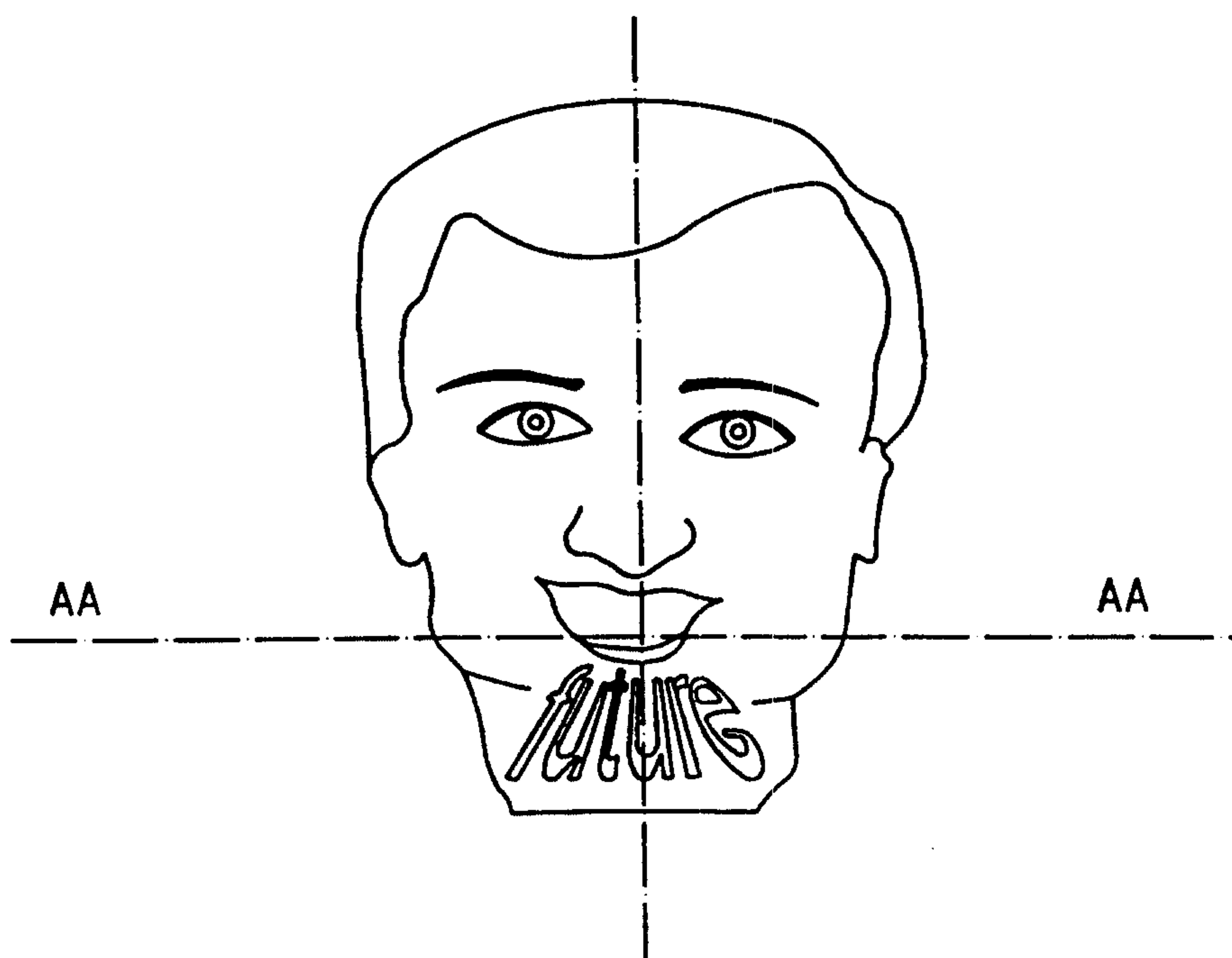


FIG. 25

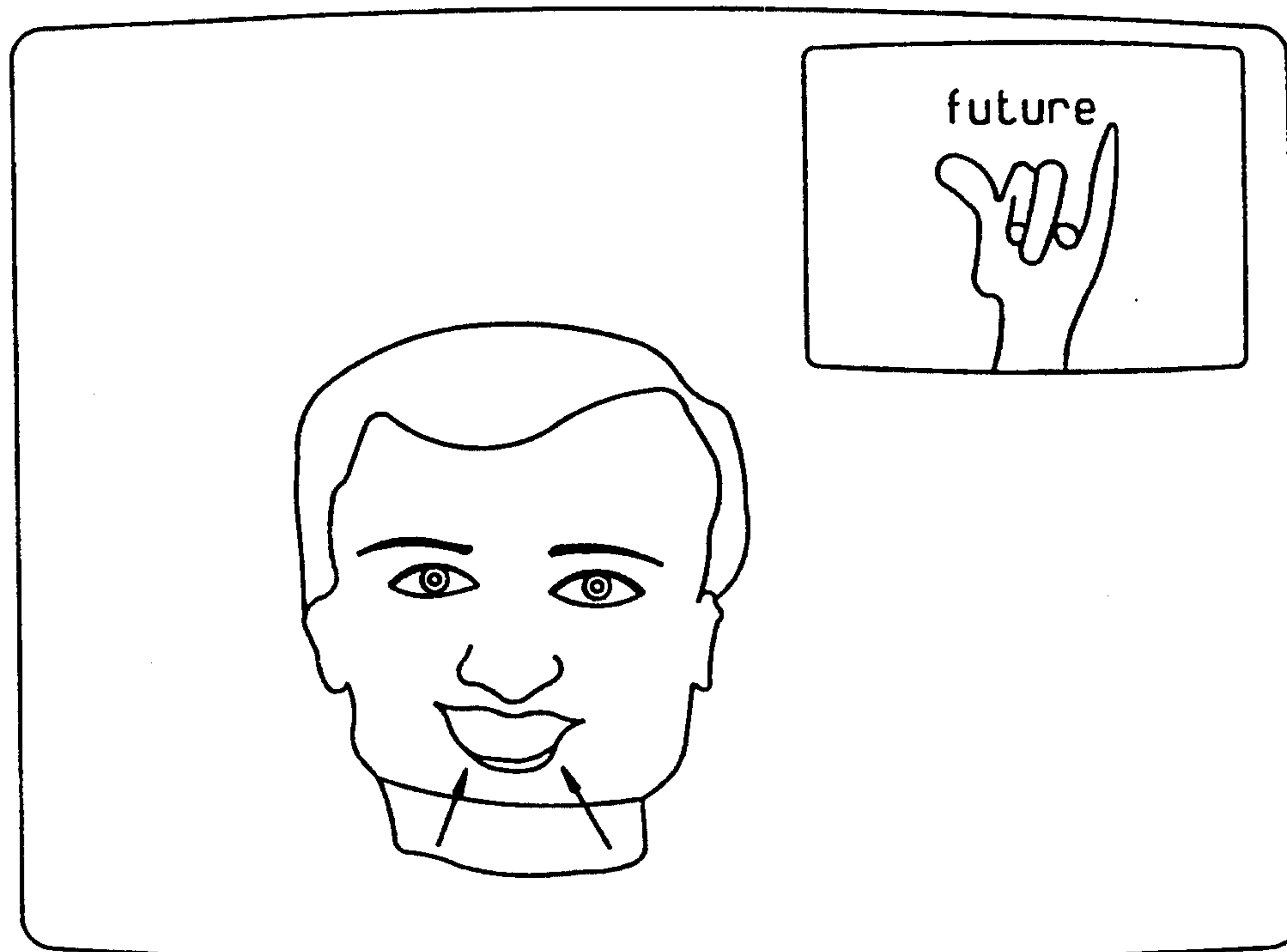


FIG. 26



FIG. 27

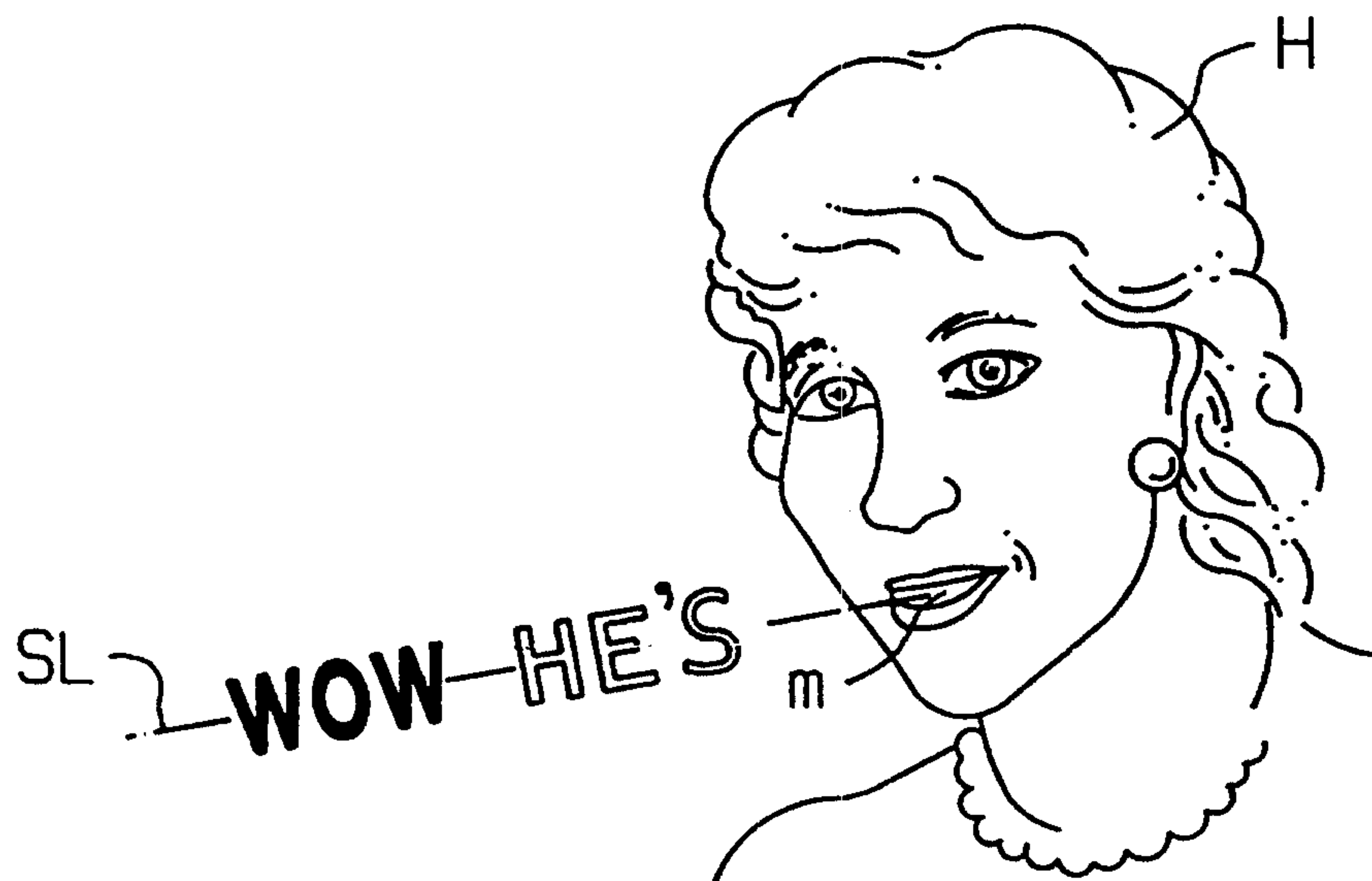


FIG. 28



FIG. 29



FIG. 30

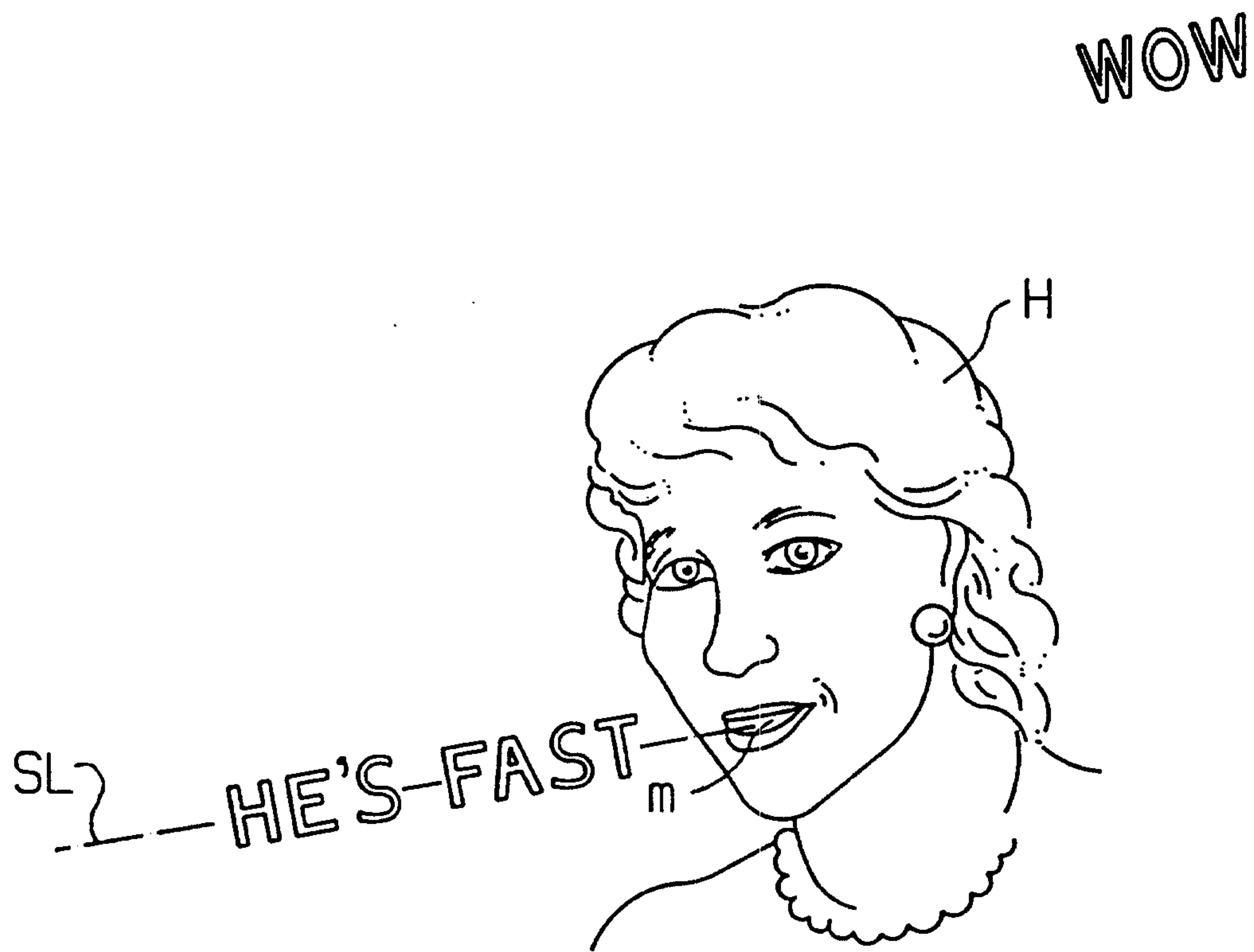


FIG. 31

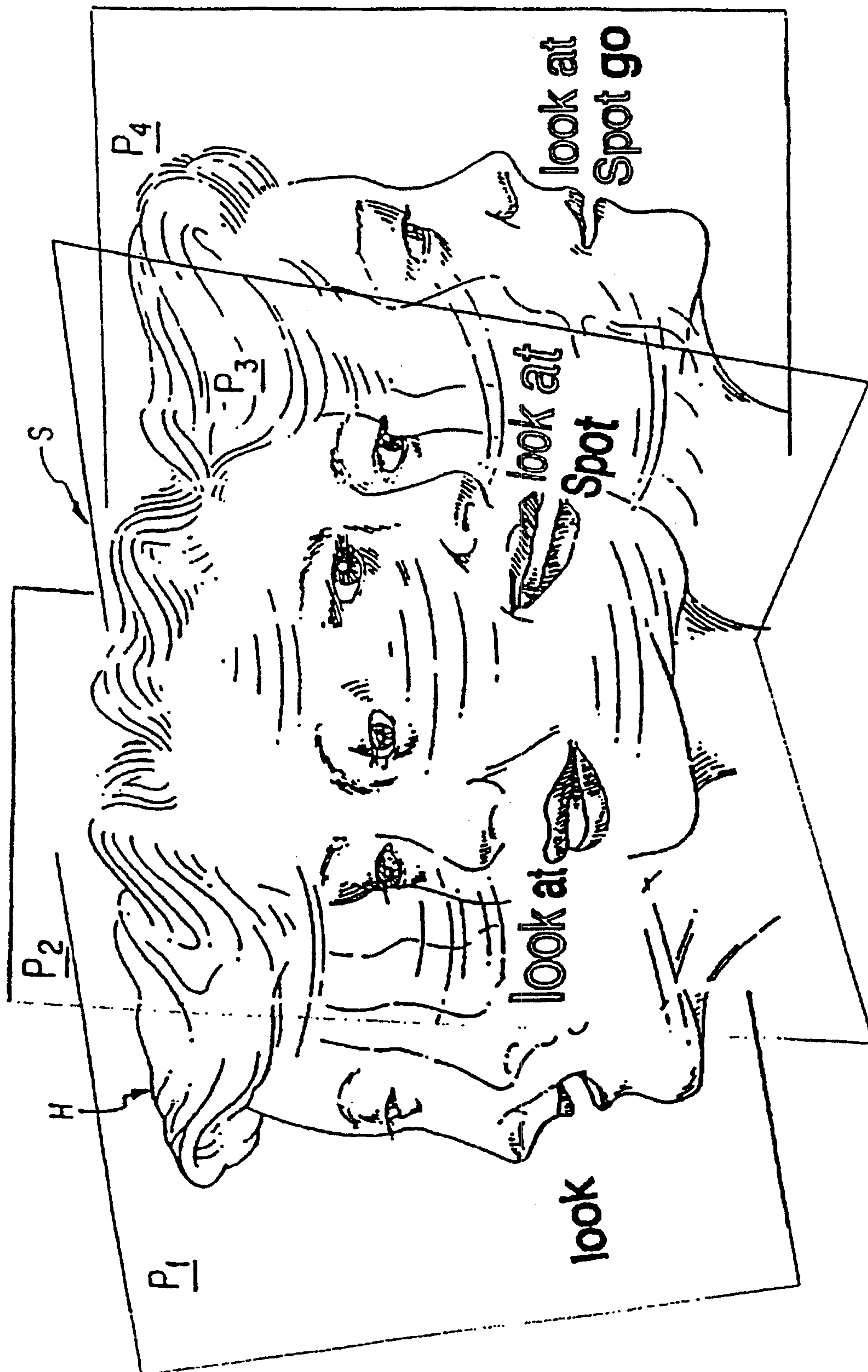


FIG. 32

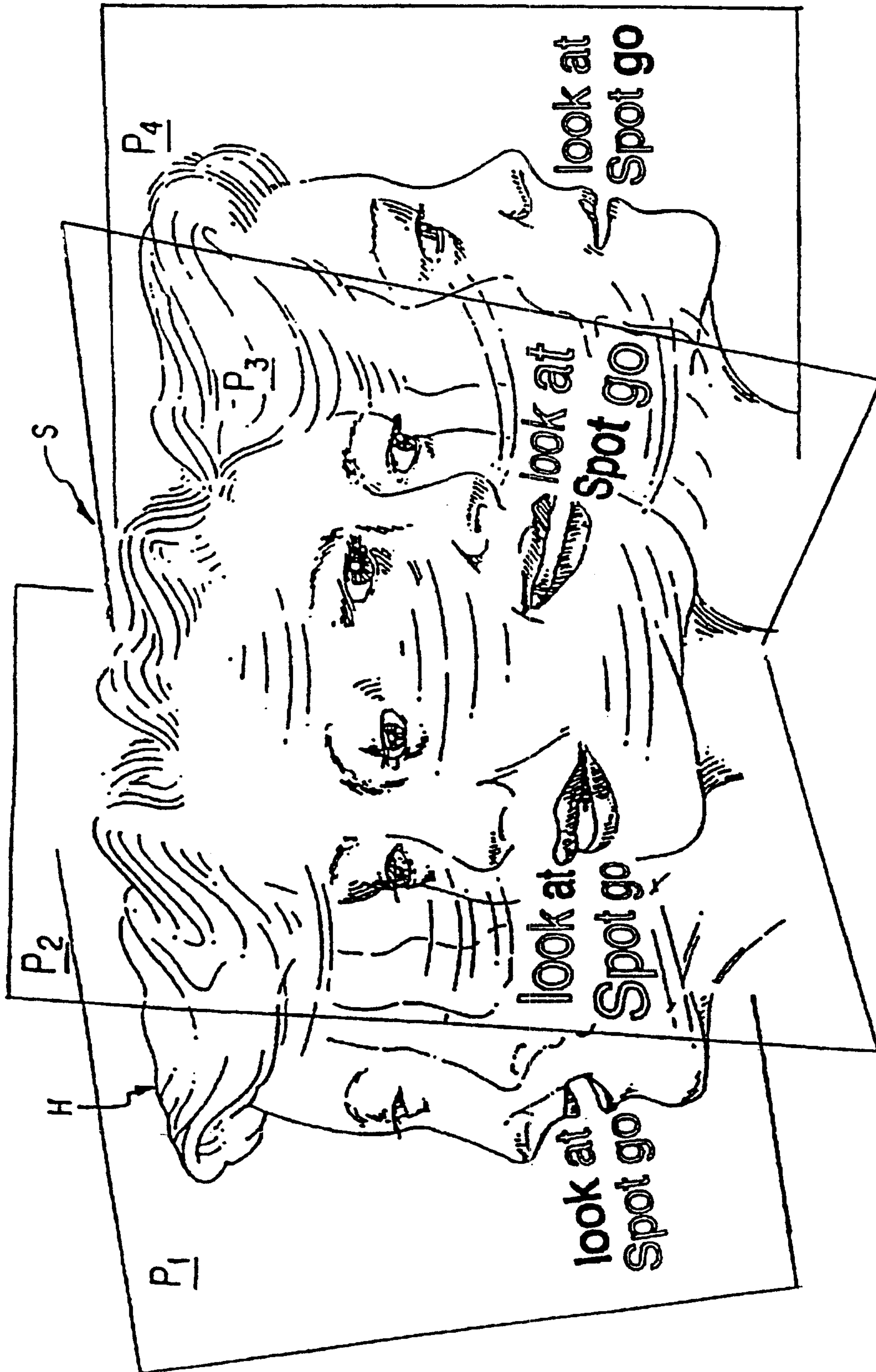


FIG. 33

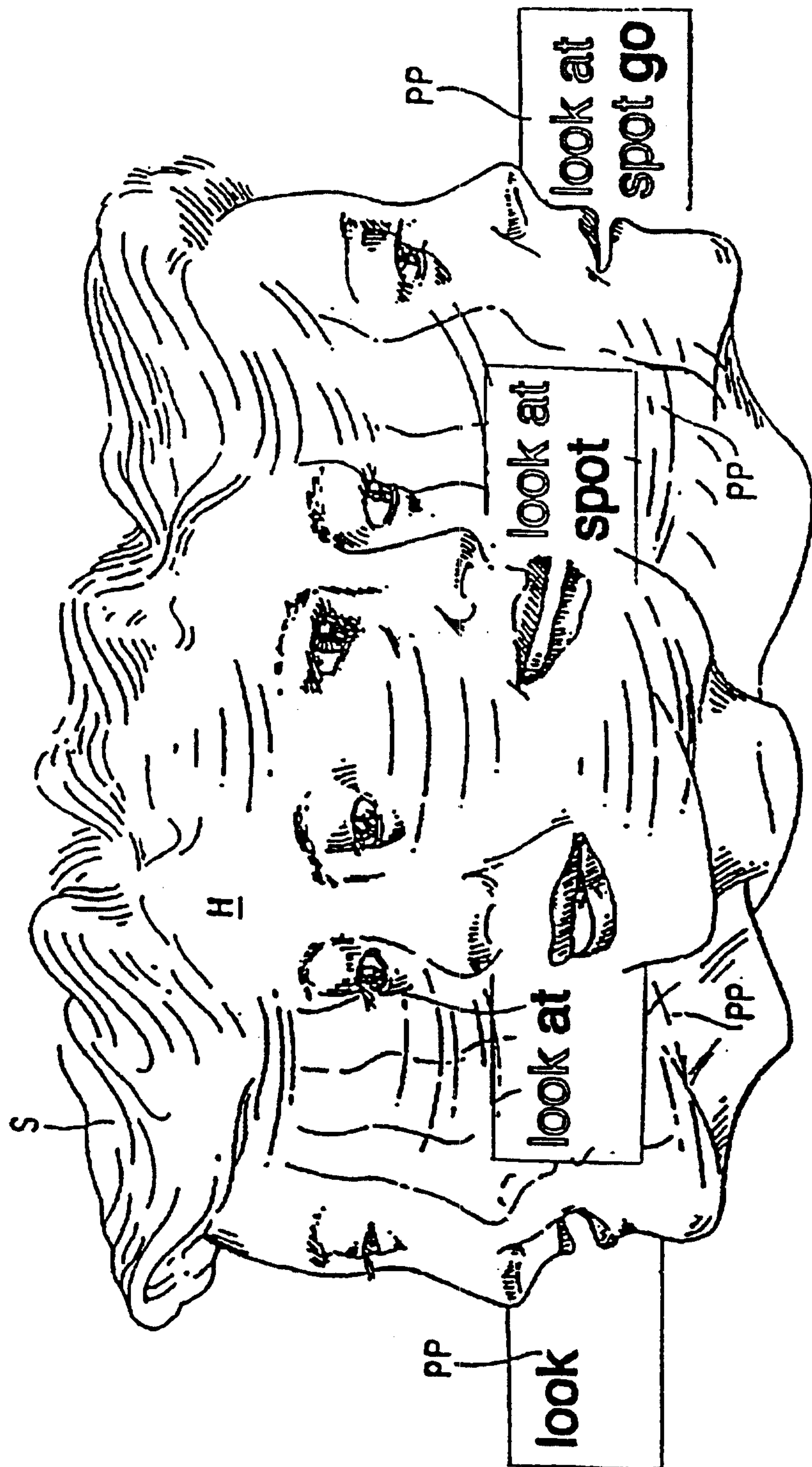


FIG. 34

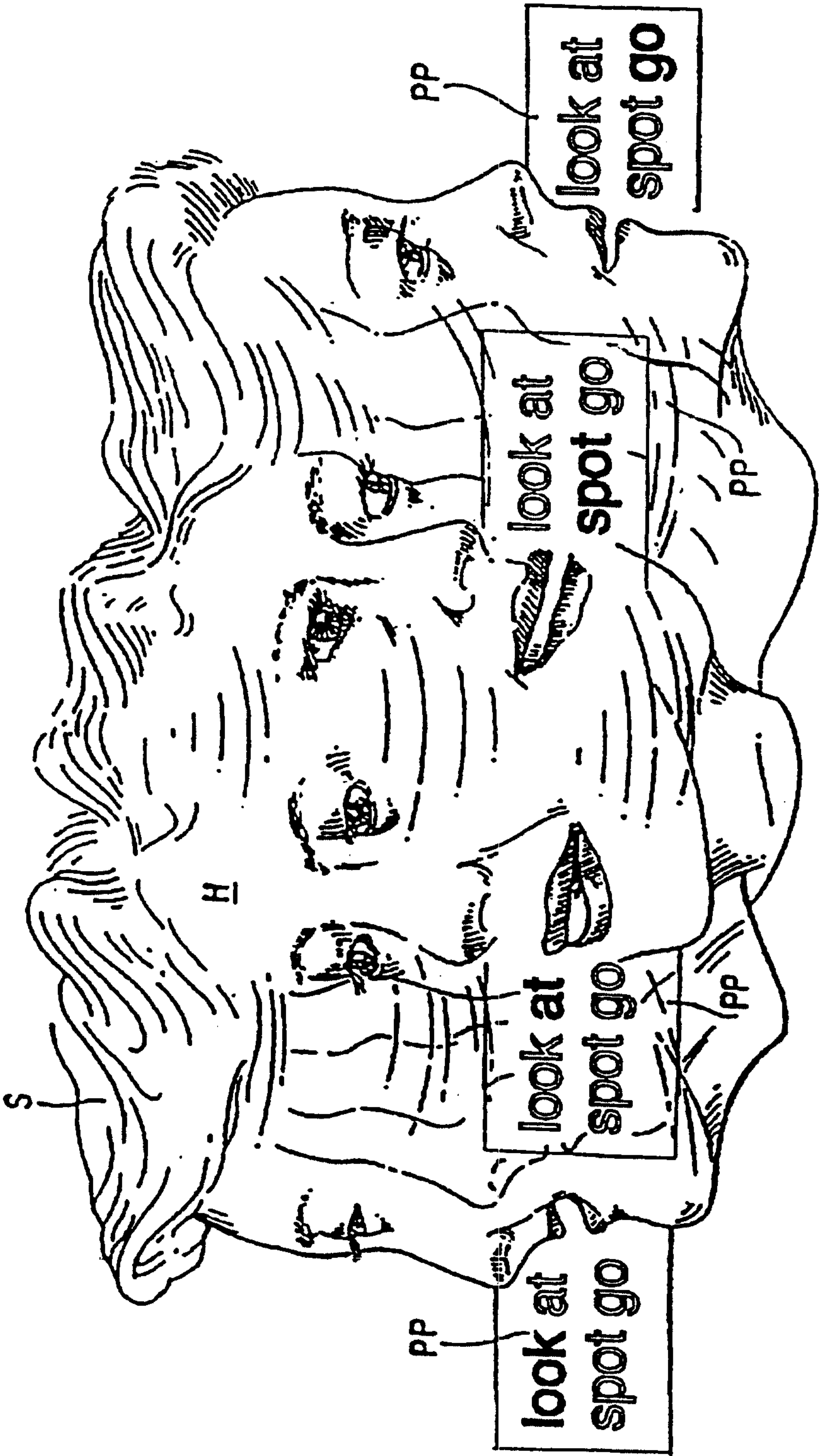


FIG. 35

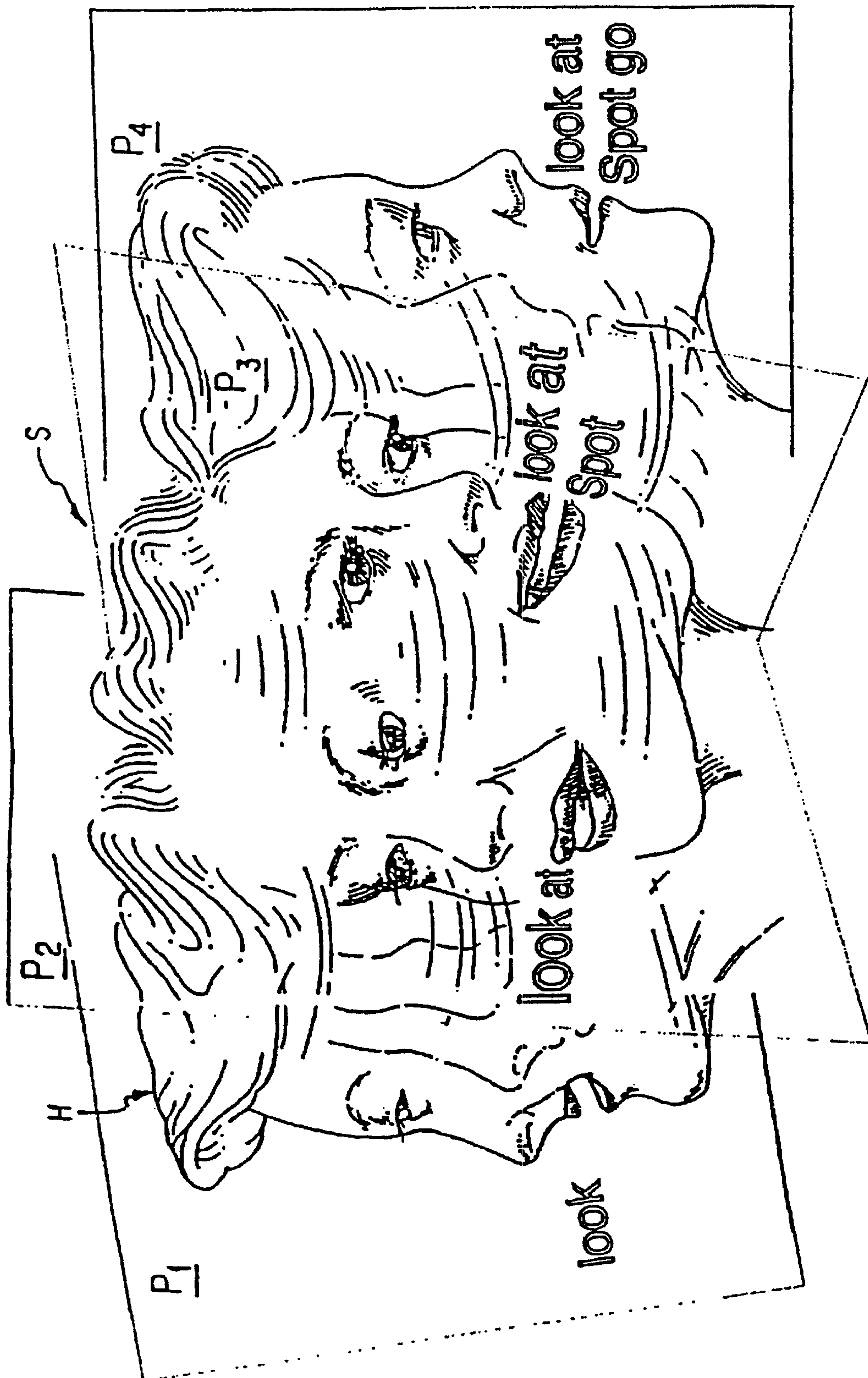


FIG. 36

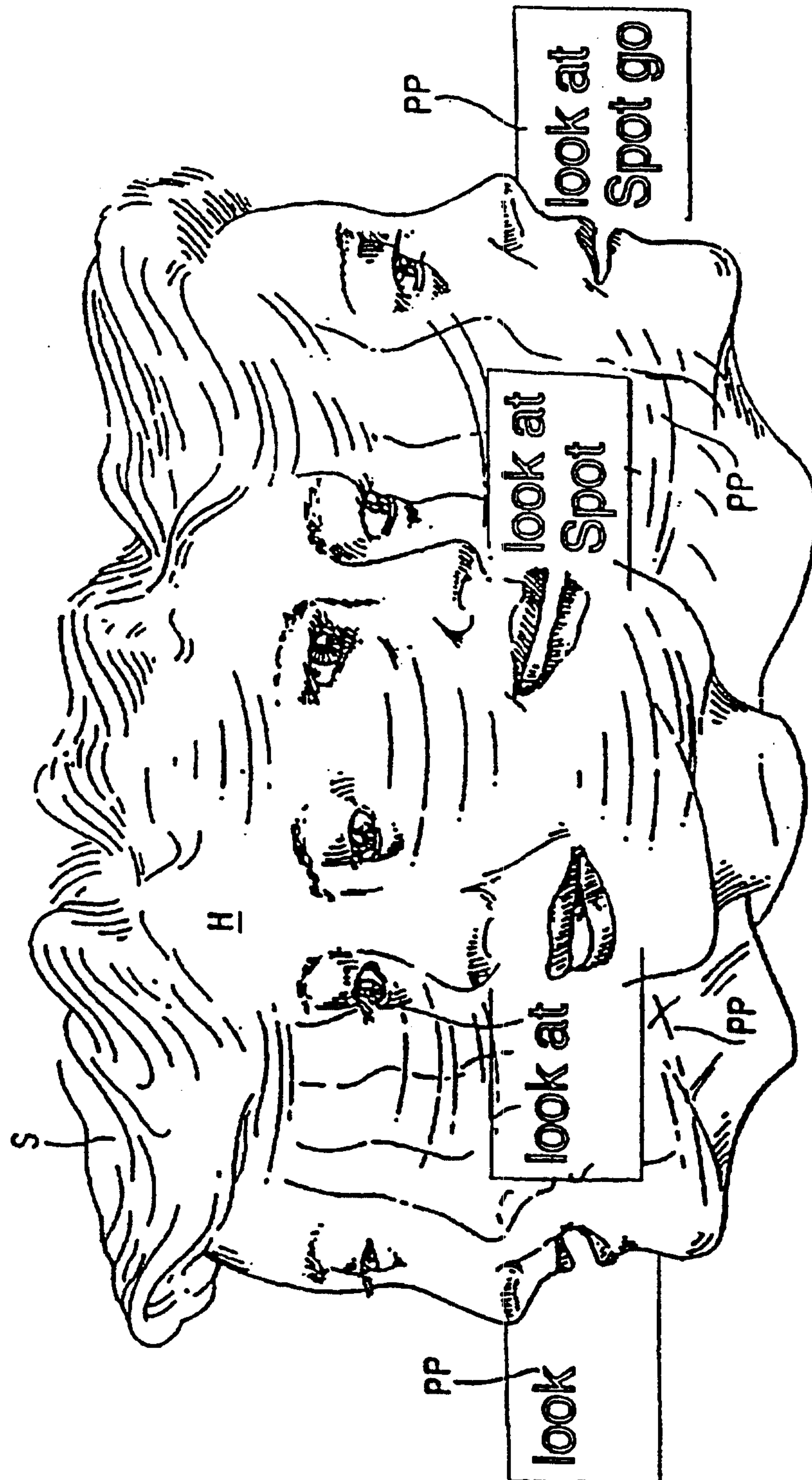


FIG. 37

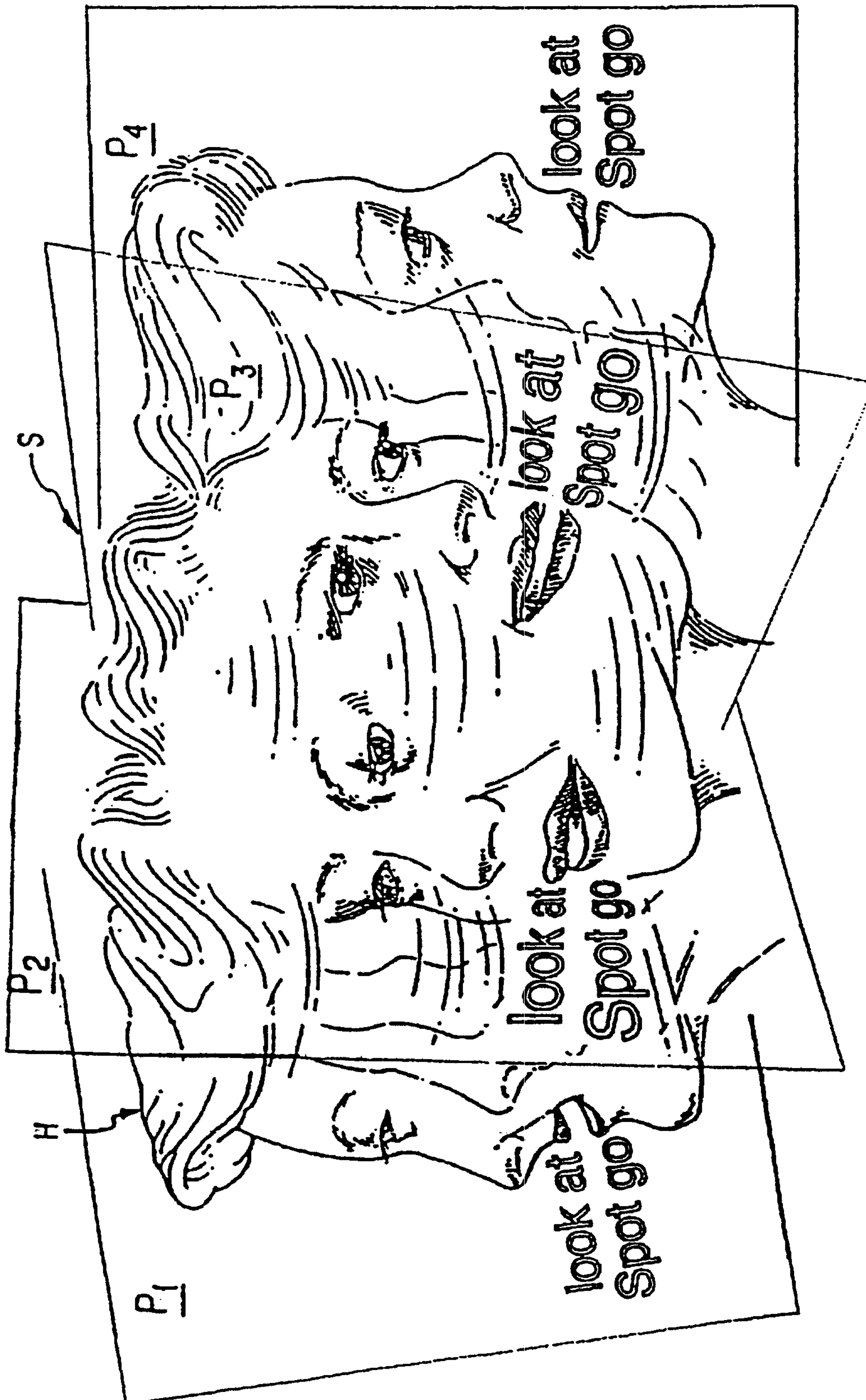


FIG. 38

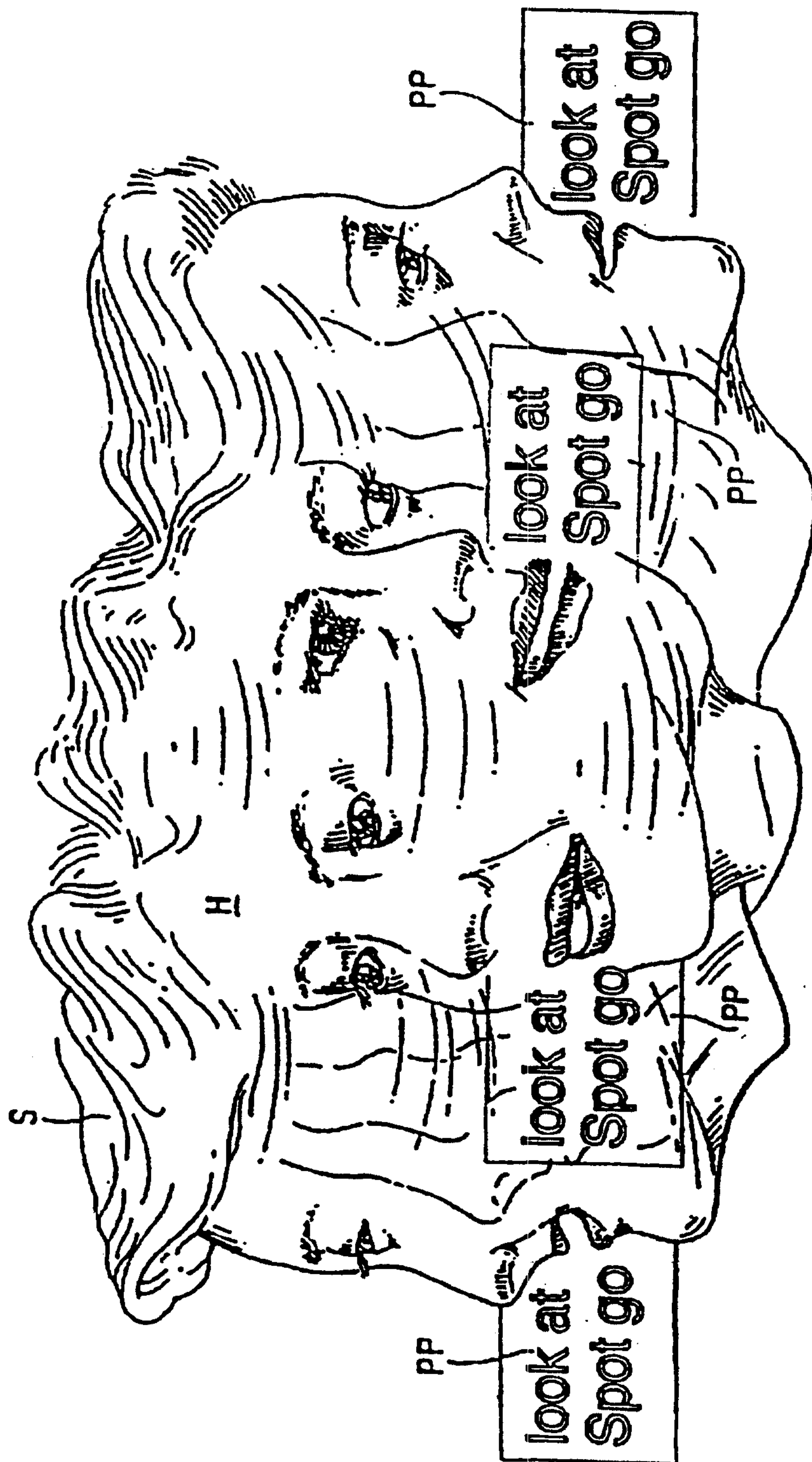


FIG. 39