

# **EXHIBIT B**



(12) **United States Patent**  
**Nir**

(10) **Patent No.:** **US 7,254,531 B2**  
(45) **Date of Patent:** **Aug. 7, 2007**

(54) **IN-CONTEXT ANALYSIS AND AUTOMATIC TRANSLATION**  
(76) Inventor: **Einat H. Nir**, 20 Komemiut St., 48039 Rosh Ha'ayin (IL)

6,446,081 B1 9/2002 Preston  
6,519,631 B1 2/2003 Rosenchein et al.  
6,609,091 B1 8/2003 Budzinski  
6,651,220 B1 11/2003 Penteroudakis et al.  
6,704,699 B2\* 3/2004 Nir ..... 704/2

(\* ) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 684 days.

\* cited by examiner  
*Primary Examiner*—Susan McFadden

(21) Appl. No.: **10/750,907**  
(22) Filed: **Jan. 5, 2004**

(65) **Prior Publication Data**  
US 2004/0138872 A1 Jul. 15, 2004

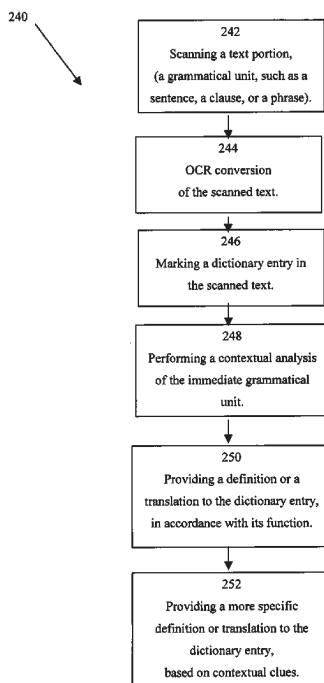
**Related U.S. Application Data**  
(63) Continuation-in-part of application No. 09/946,391, filed on Sep. 6, 2001, now Pat. No. 6,704,699.  
(60) Provisional application No. 60/229,794, filed on Sep. 5, 2000.

(51) **Int. Cl.**  
**G06F 17/20** (2006.01)  
(52) **U.S. Cl.** ..... **704/10; 382/313**  
(58) **Field of Classification Search** ..... **704/10; 382/313**  
See application file for complete search history.

(56) **References Cited**  
U.S. PATENT DOCUMENTS  
5,680,511 A \* 10/1997 Baker et al. .... 704/257

(57) **ABSTRACT**  
Methods are described for automatically presenting at least one definition in context, for a dictionary entry, based a contextual analysis of the grammatical unit, such as the sentence, the clause, or the phrase, in which the dictionary entry is embedded. The dictionary entry may be single word, for example, "consideration," or a word combination, such as "due to," provided the word combination appears in an ordinary dictionary as a single entry. A definition in context will generally include a definition in accordance with the function of the word in the grammatical unit, and may further include a preferred definition from among several options for the same function, based on contextual clues within the grammatical unit. Similarly, a translation in context may be presented. The present invention further relates to a method of preparing automatic-translation viable documents, in a first language, adapted for accurate automatic translation, to a plurality of languages. Additionally, interactive automatic translation software is taught, for use in tandem with a human translator. These may be performed by a supercomputer, a personal computer, a laptop, a palm-top, a stand-alone hand-held, computerized scanning apparatus, or over the web.

**9 Claims, 42 Drawing Sheets**



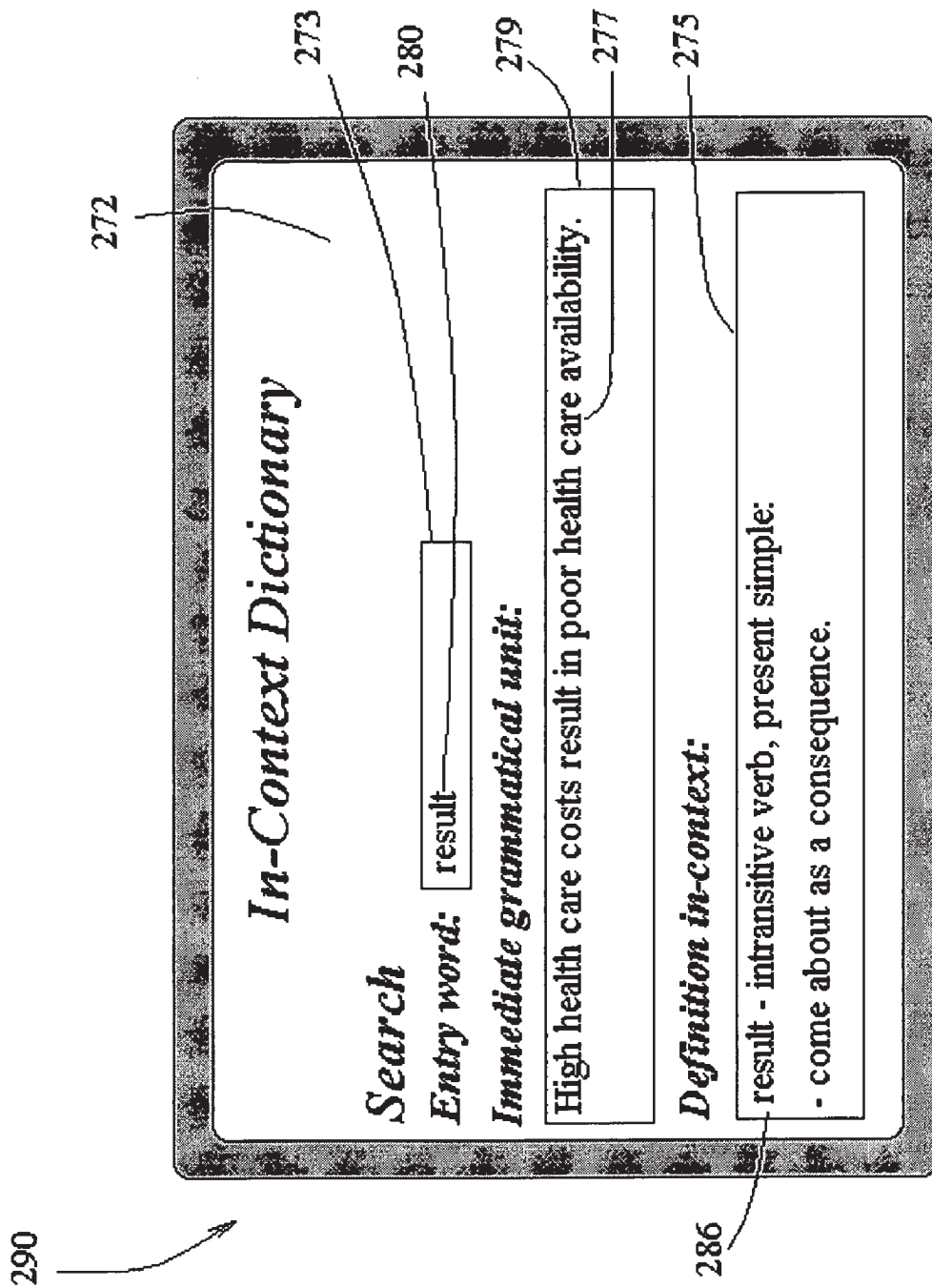


Fig. 1A

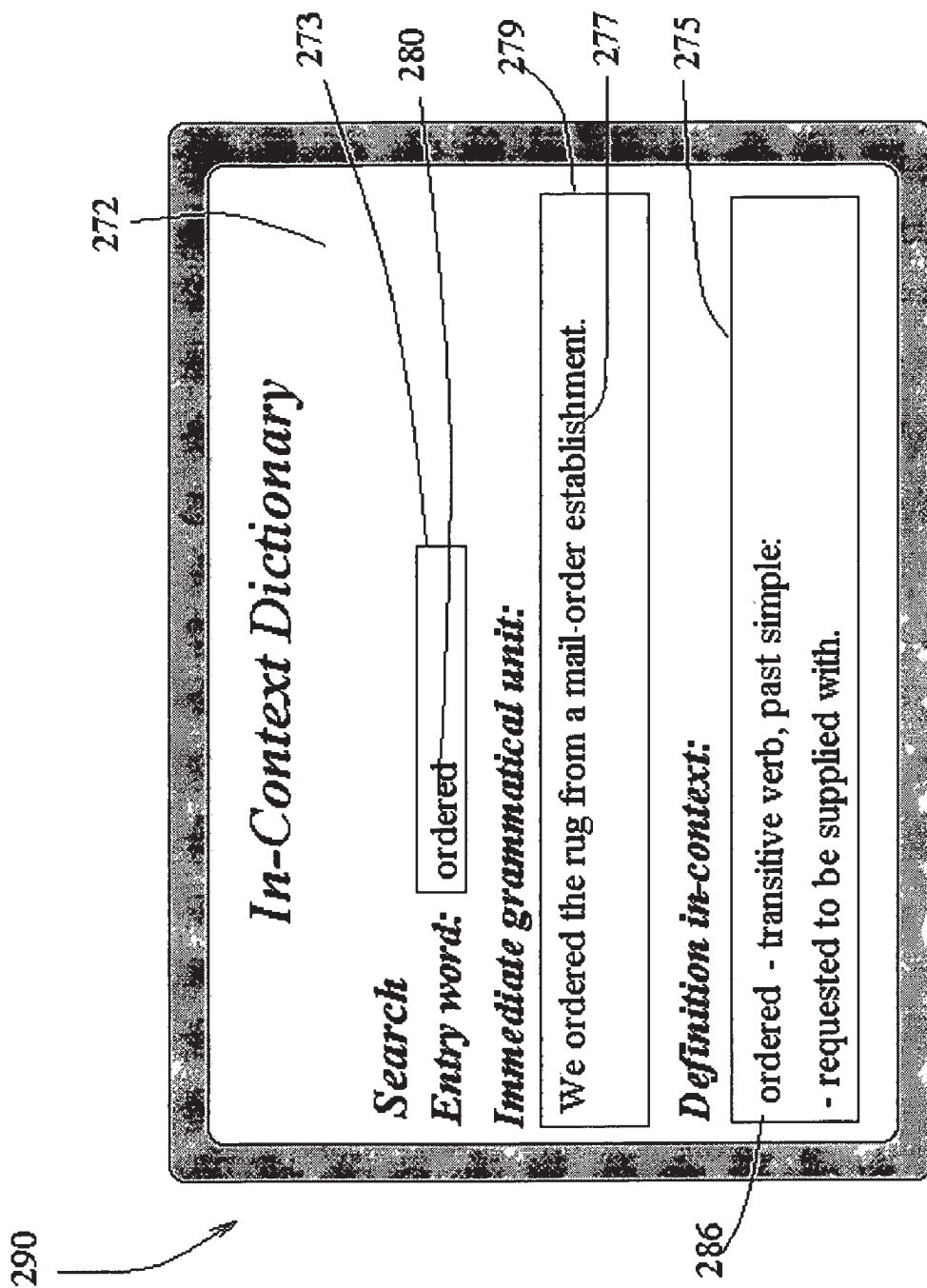
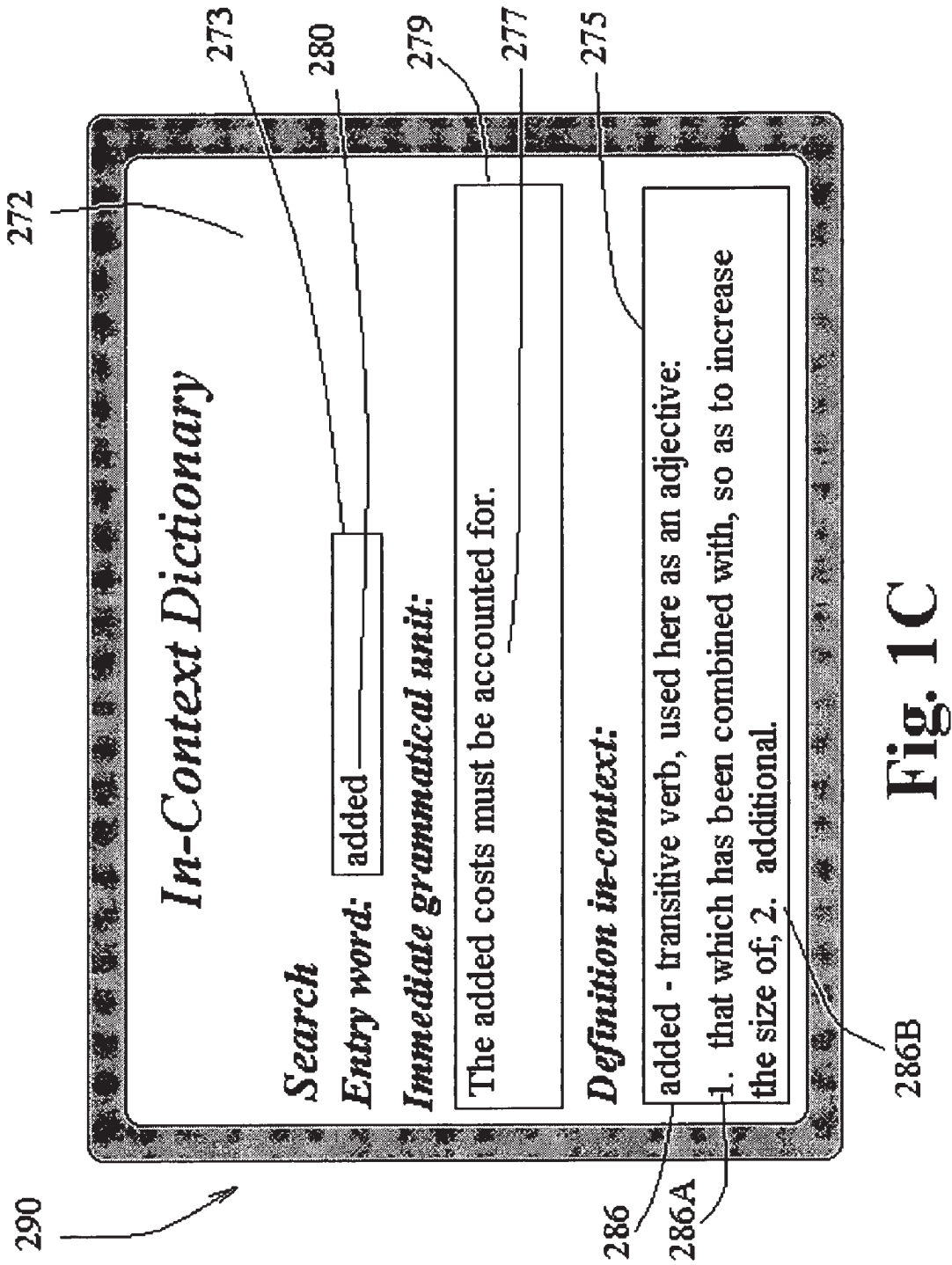


Fig. 1B



**Fig. 1C**

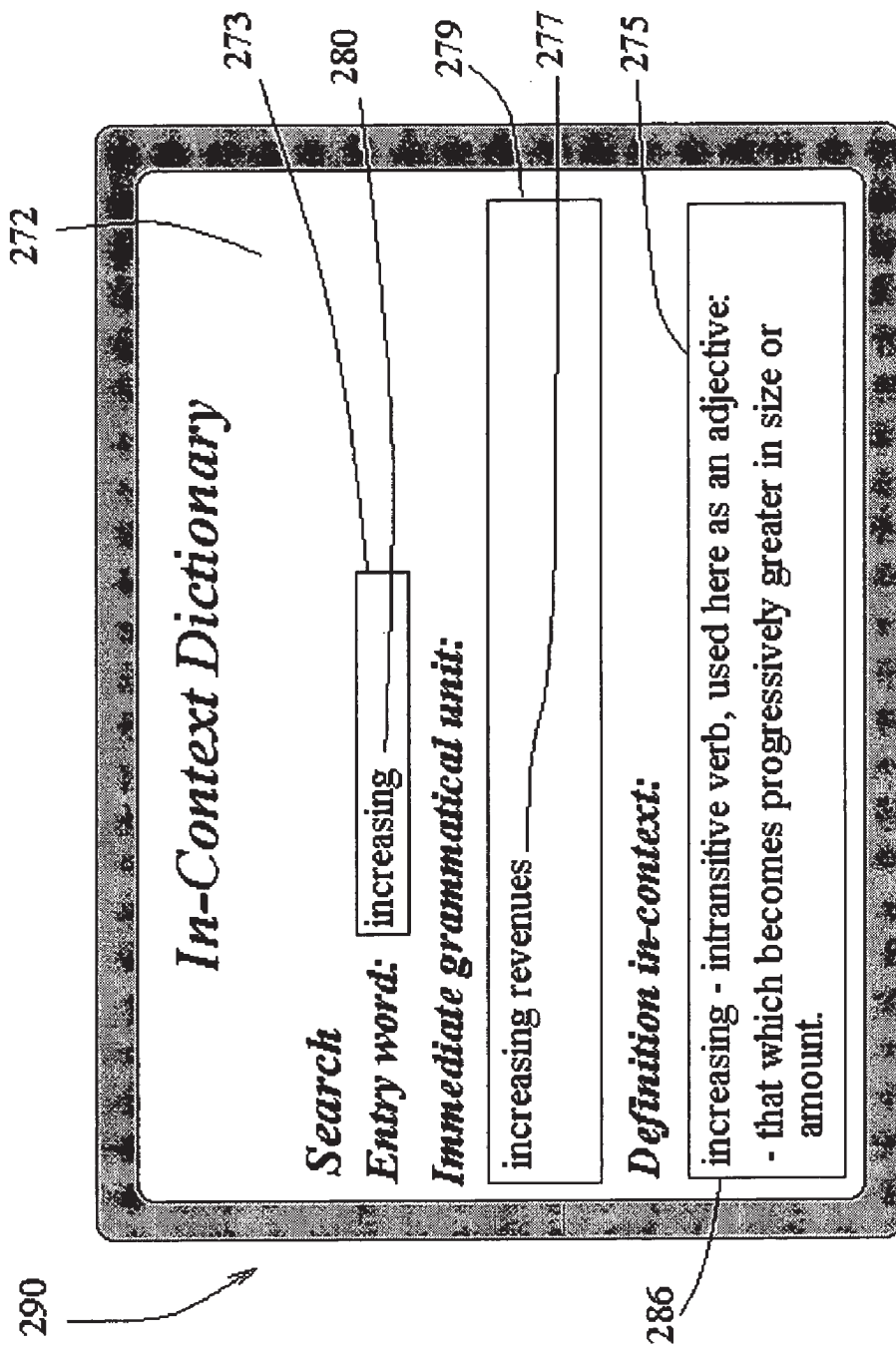


Fig. 1D

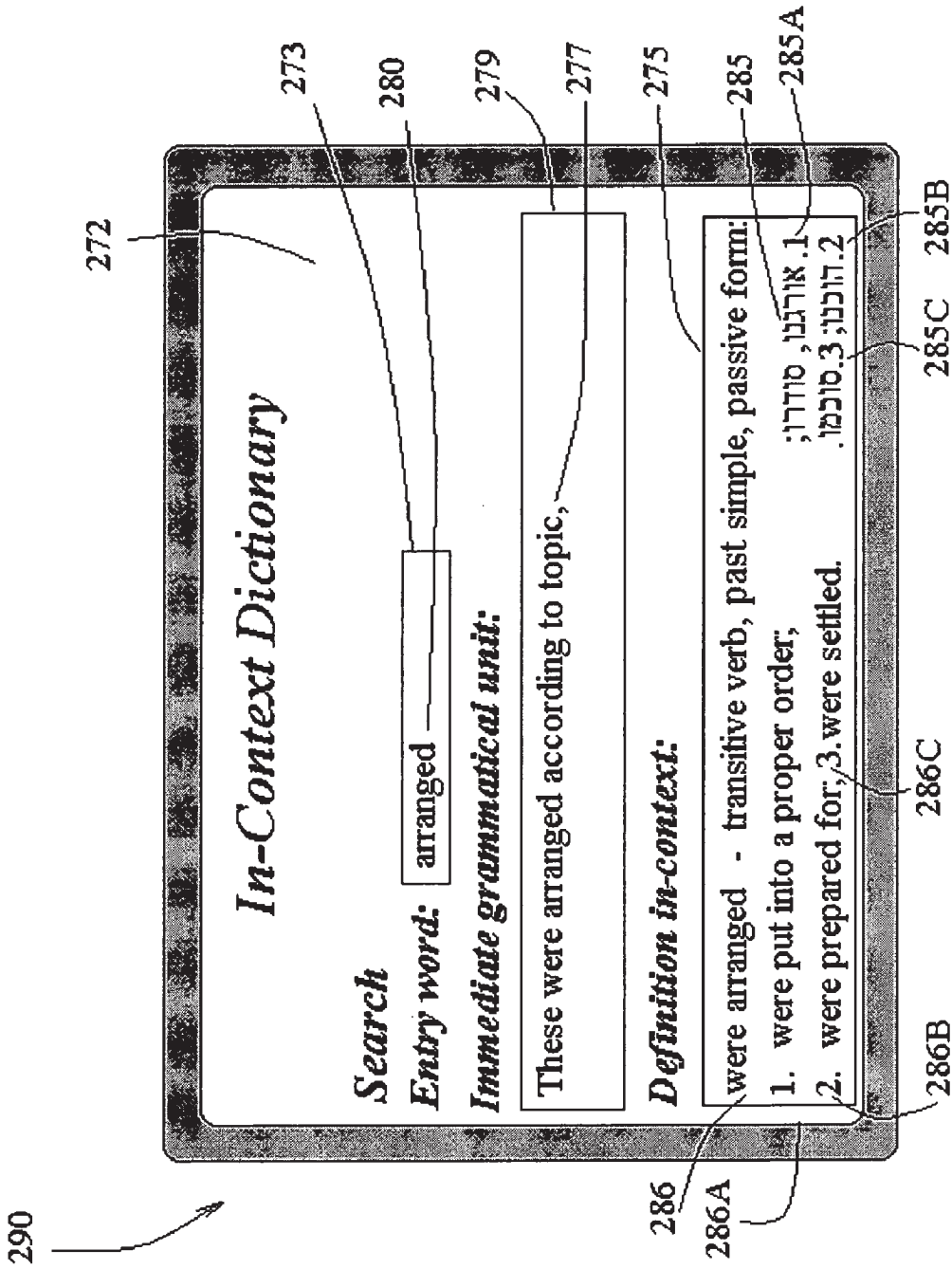


Fig. 1E

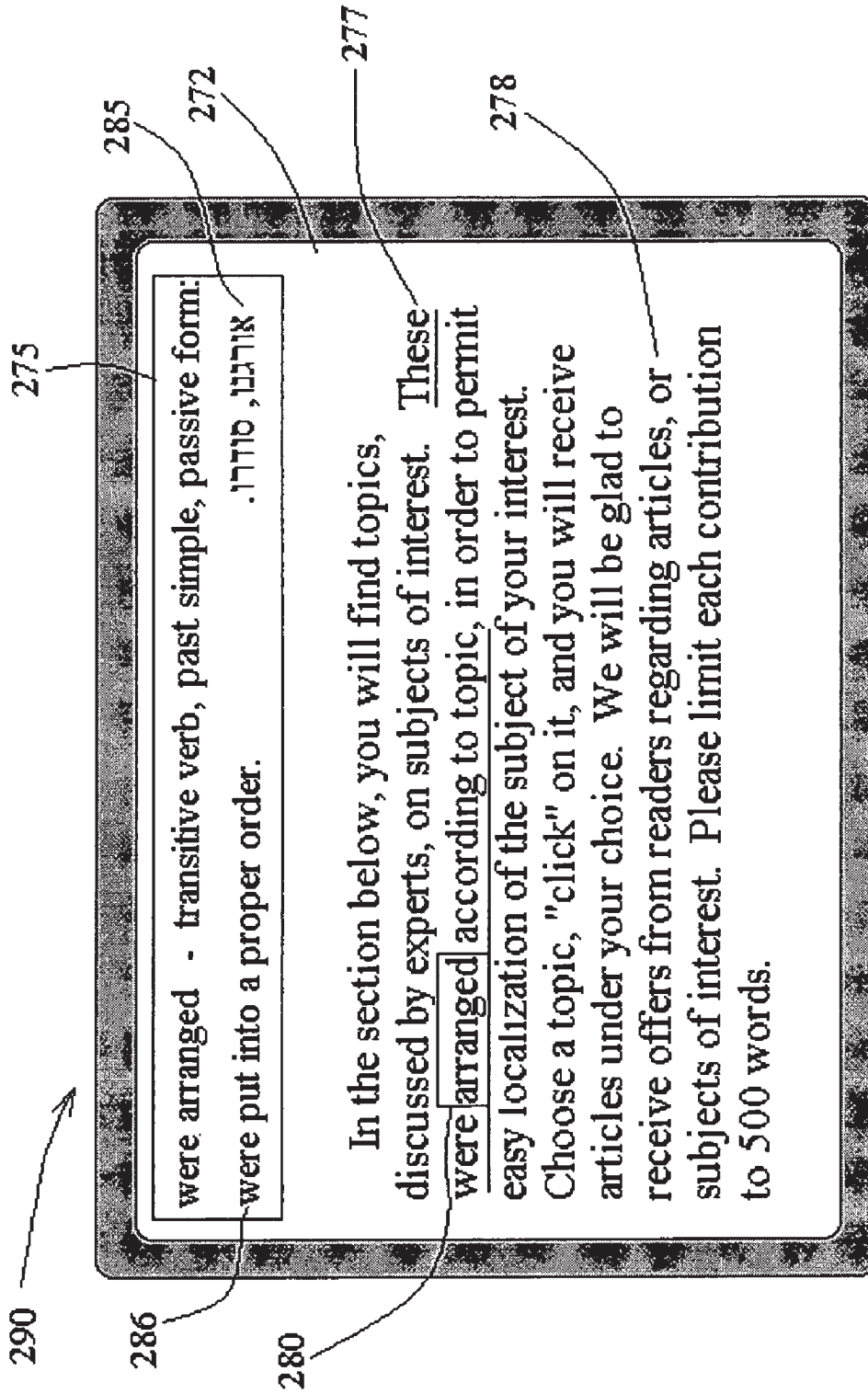
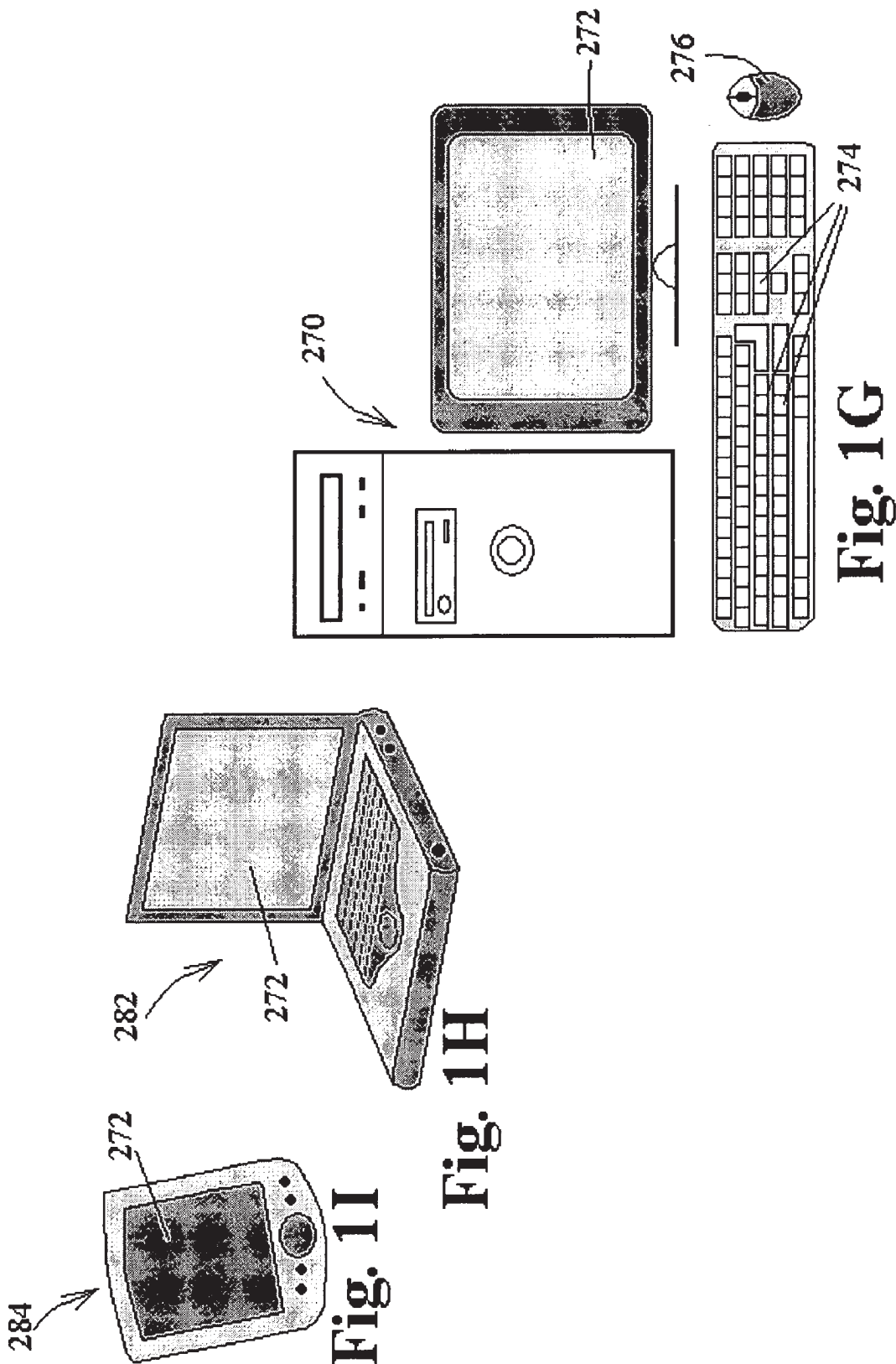


Fig. 1F





270

272

276

Fig. 1G

274

282

272

Fig. 1H

284

272

Fig. 1I

**Fig. 1J**

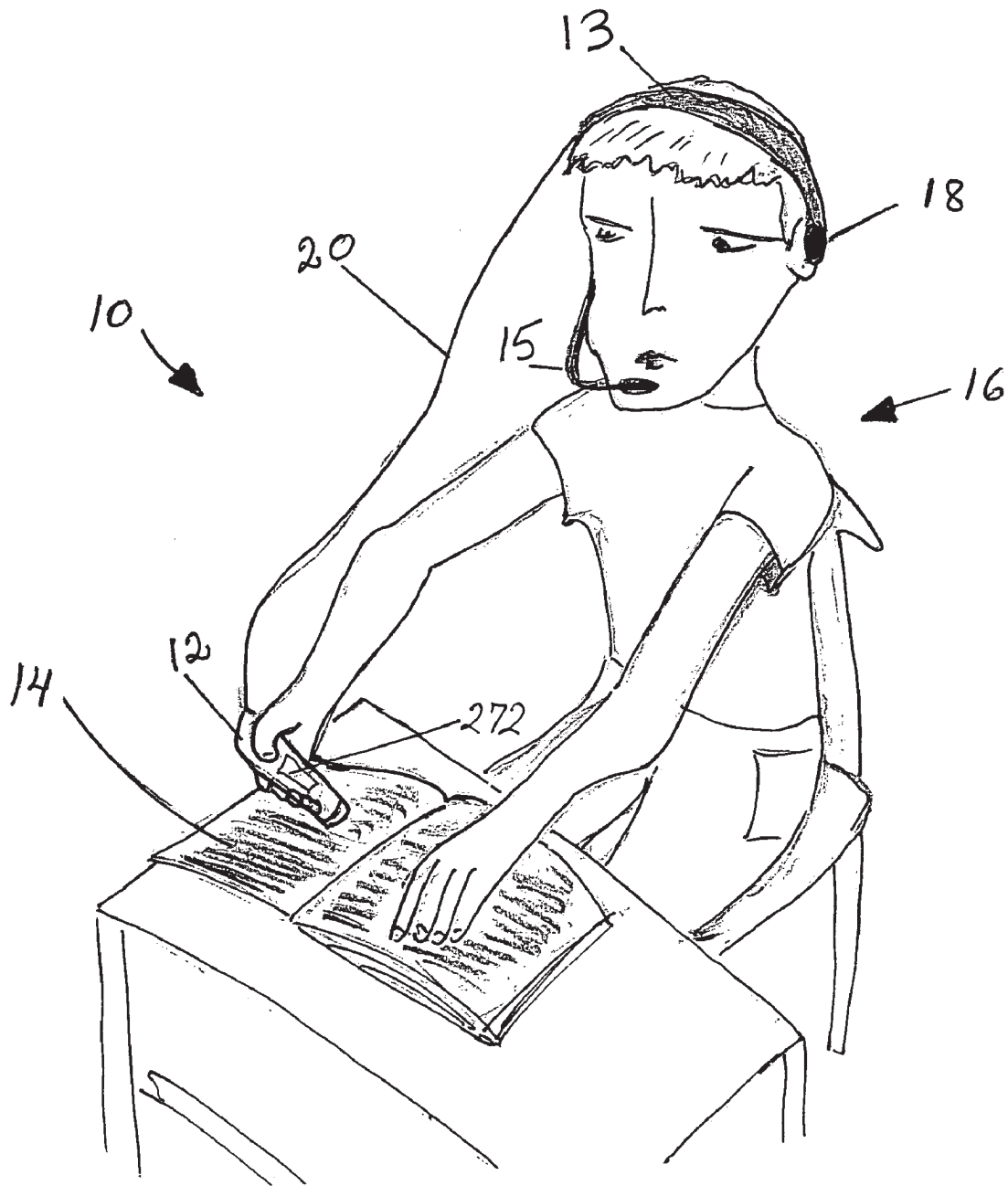


Fig. 2A

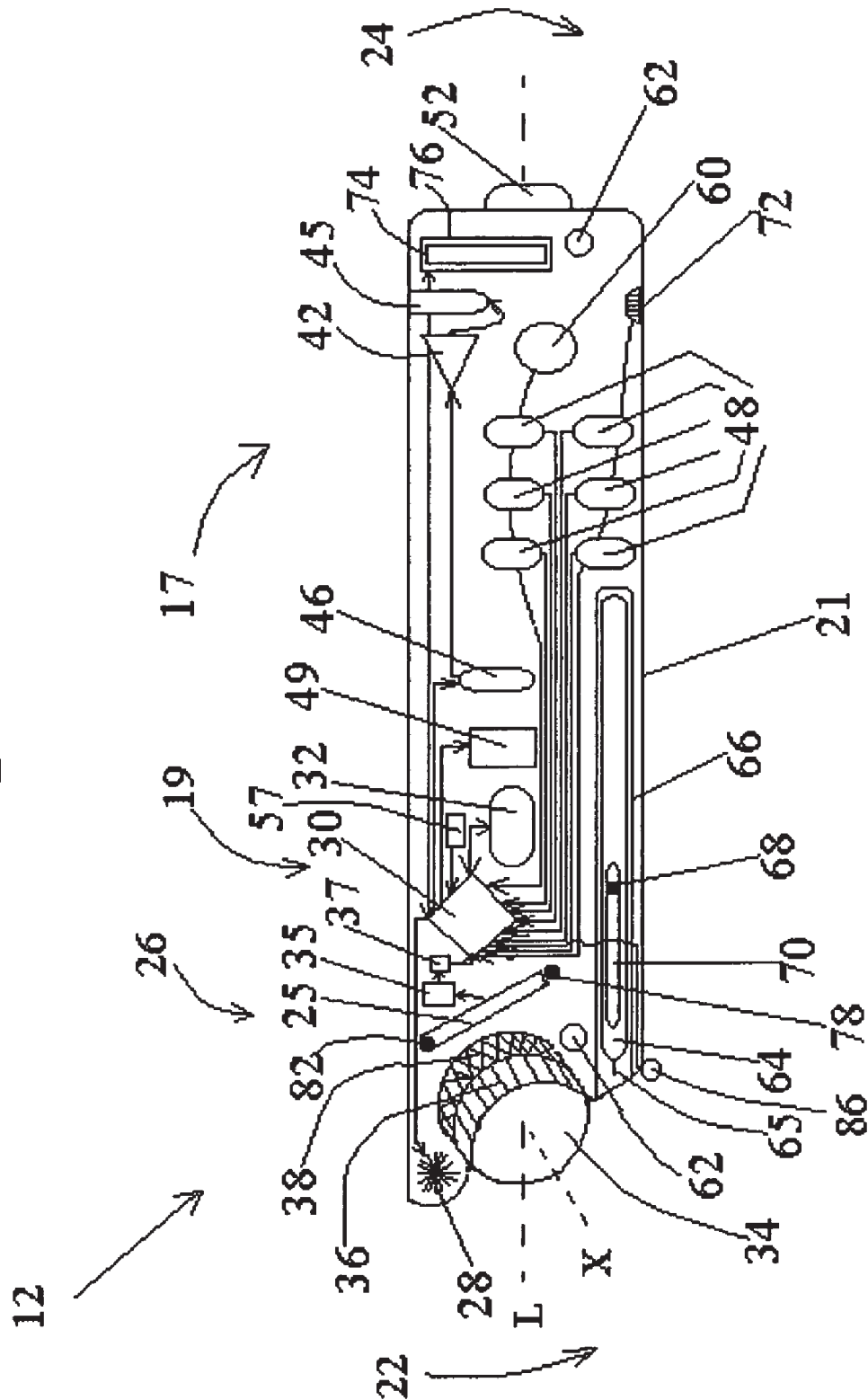
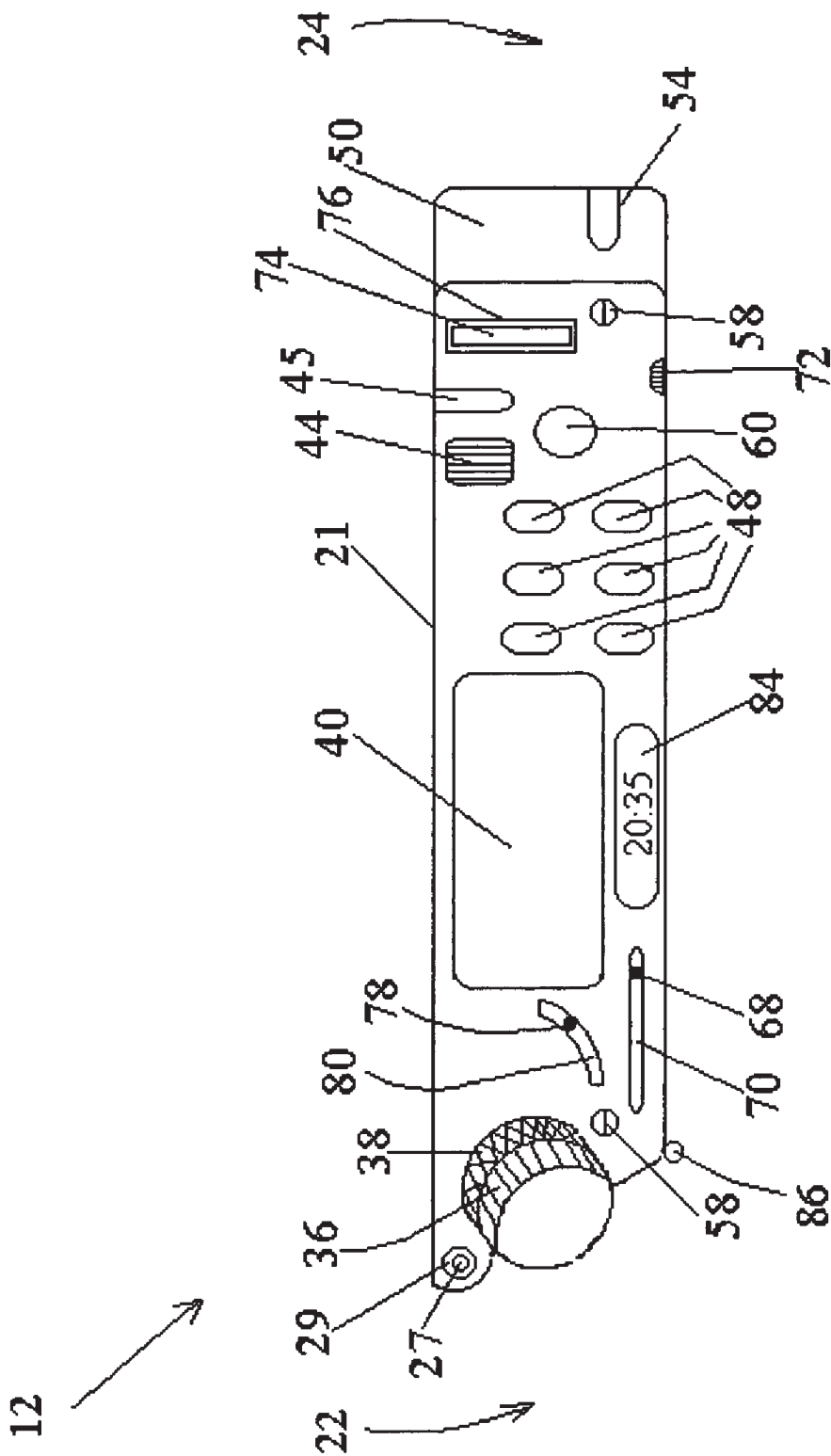
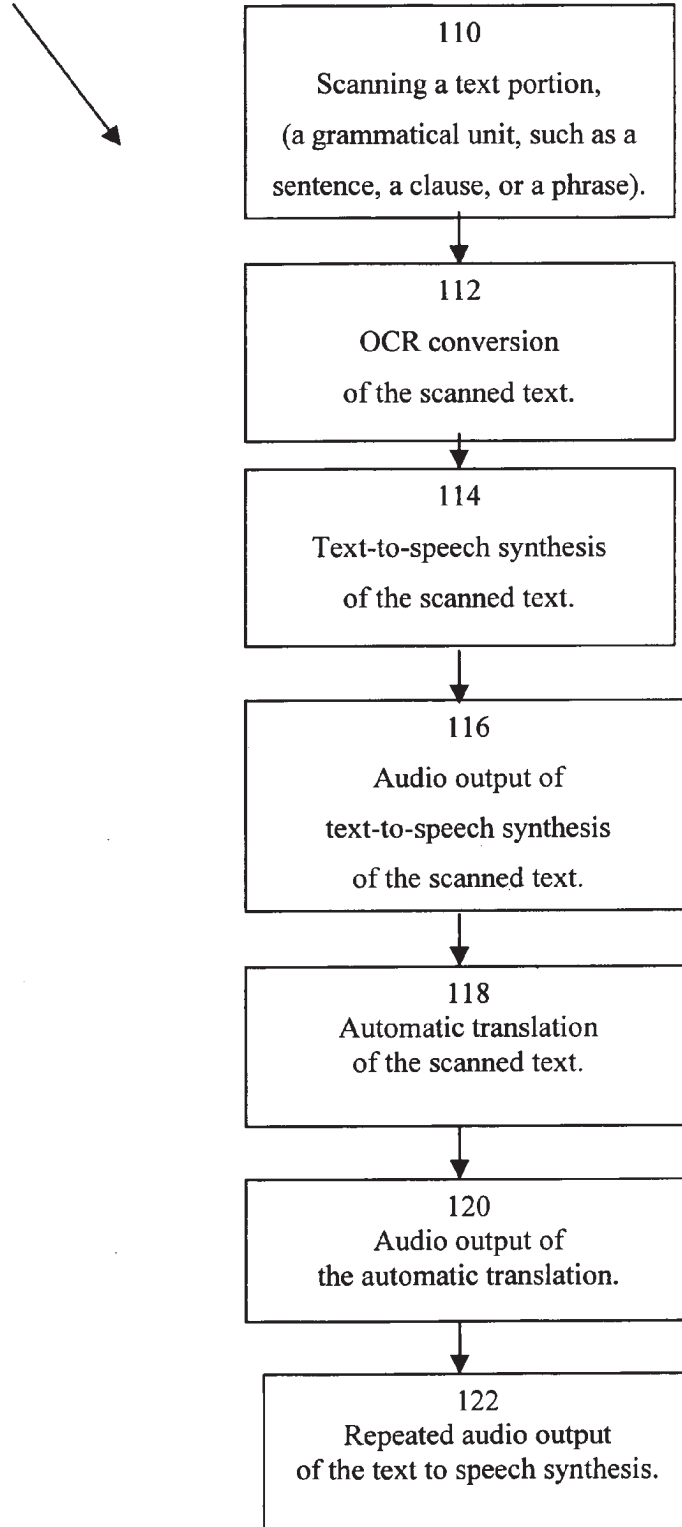


Fig. 2B



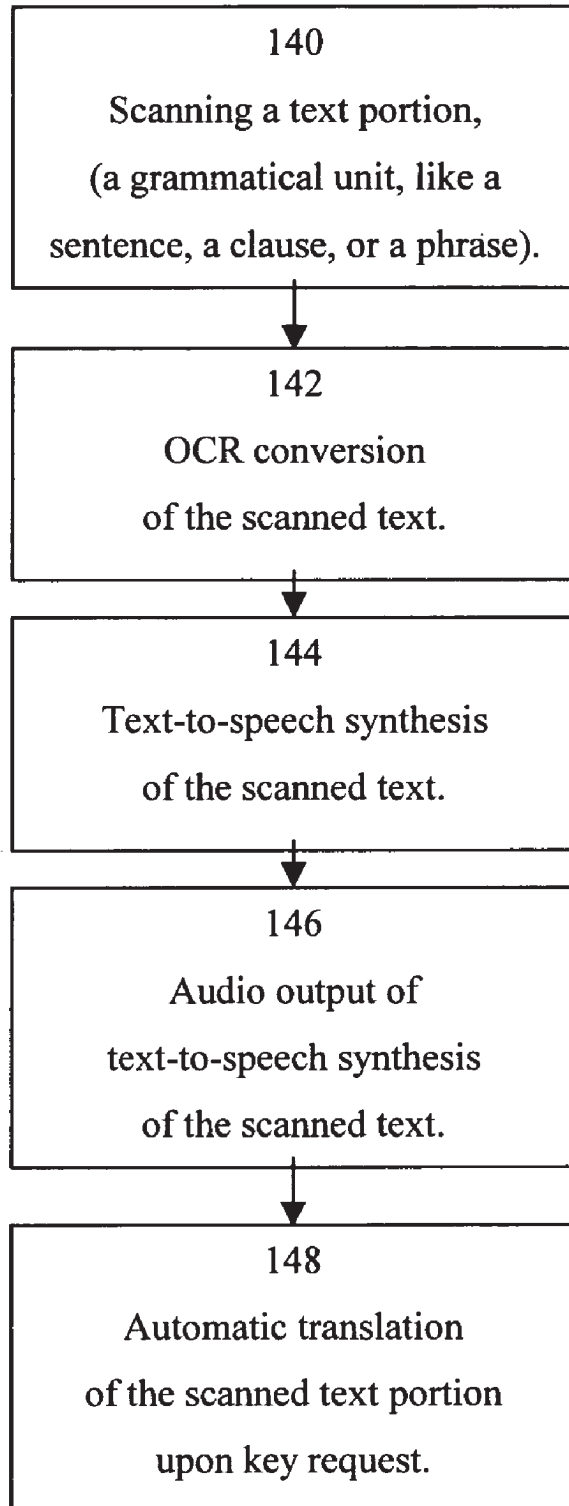
**Fig. 3**

100

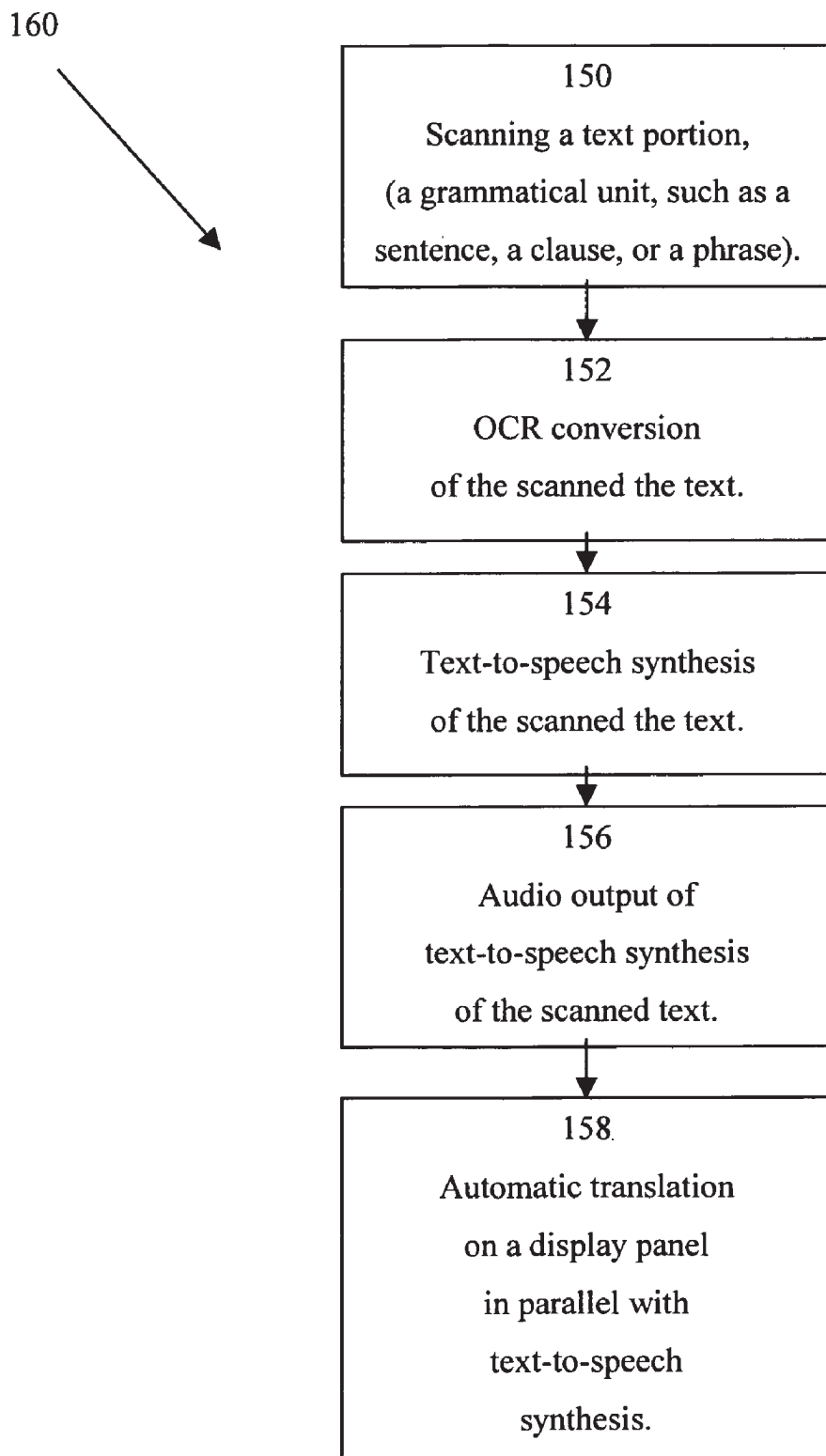


**Fig. 4**

130



**Fig. 5**



**Fig. 6A**

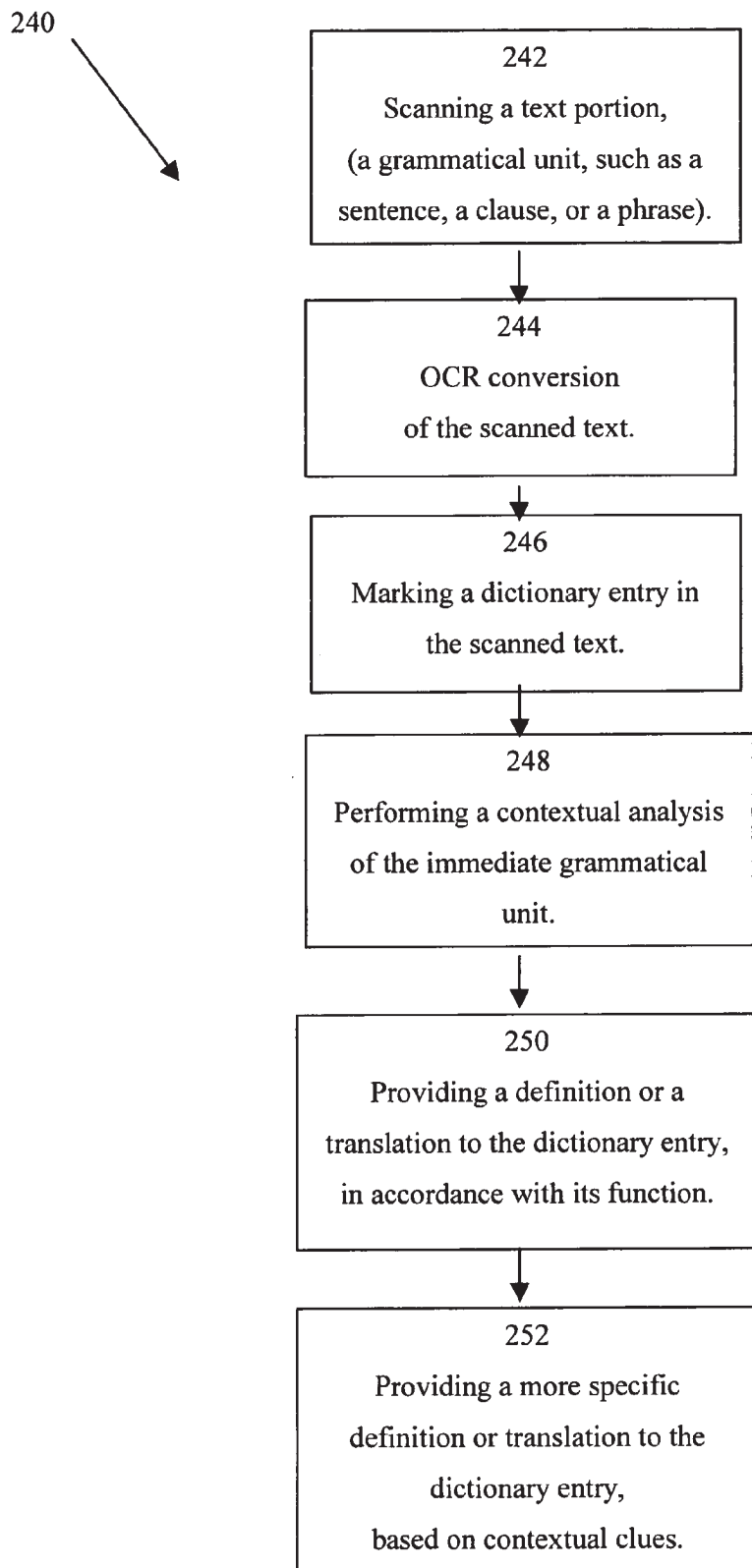




Fig. 6B

260

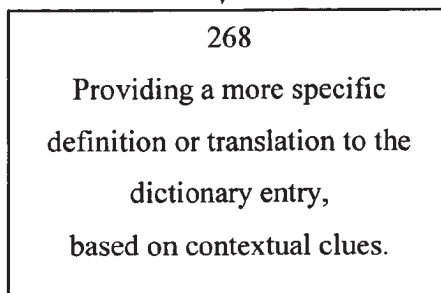
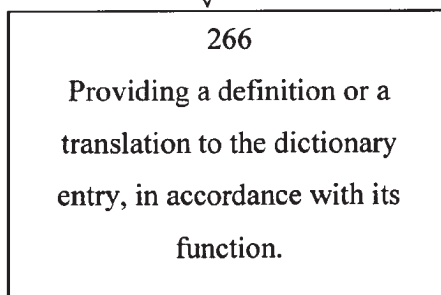
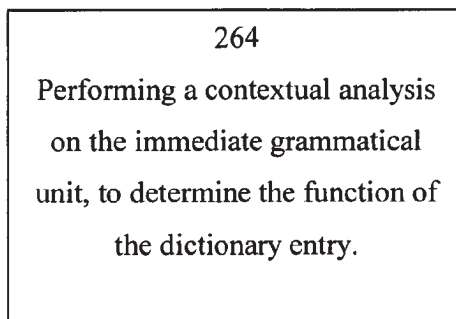
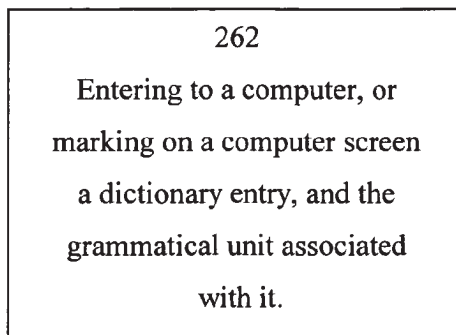
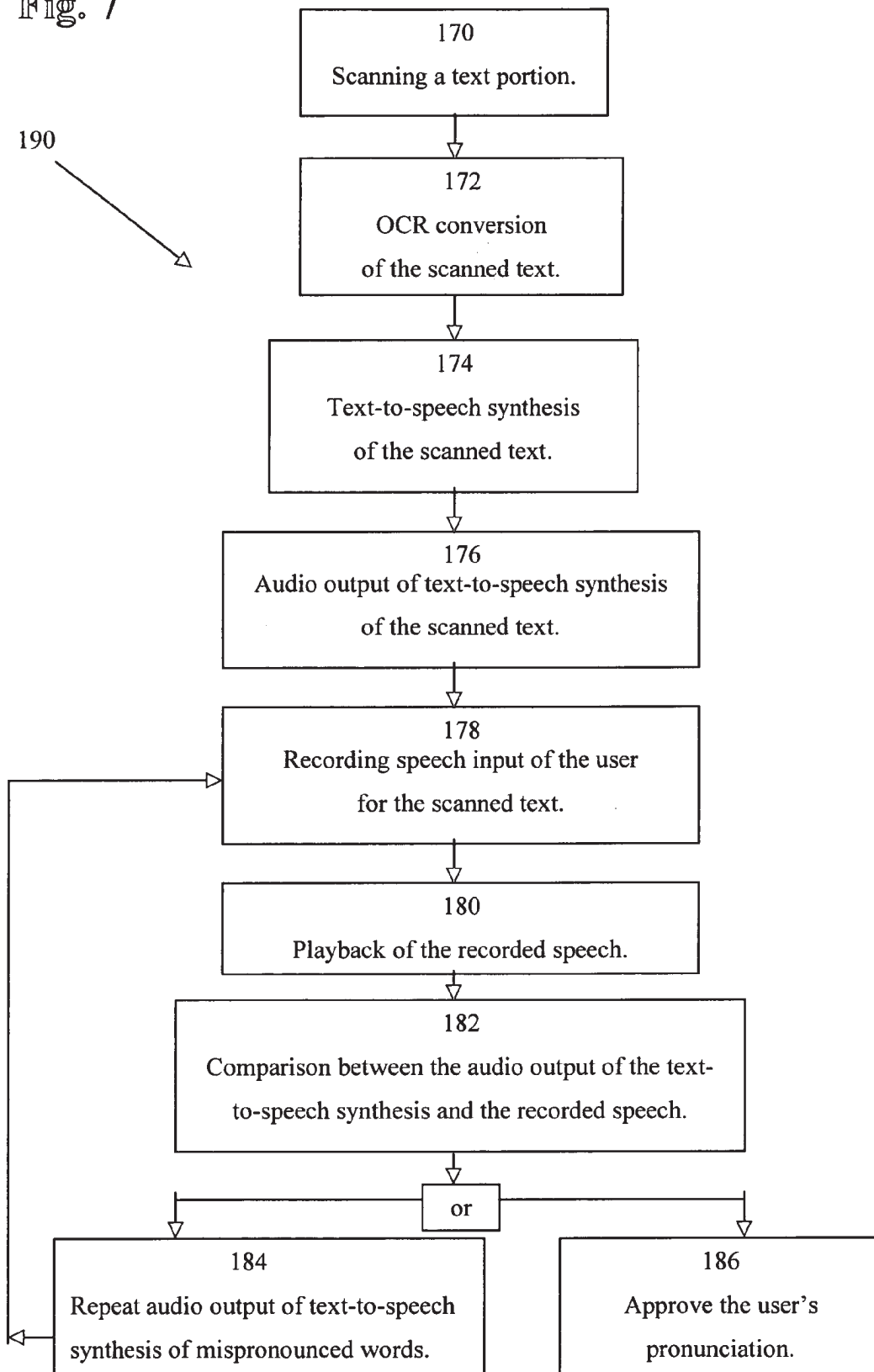


Fig. 7



**Fig. 8**

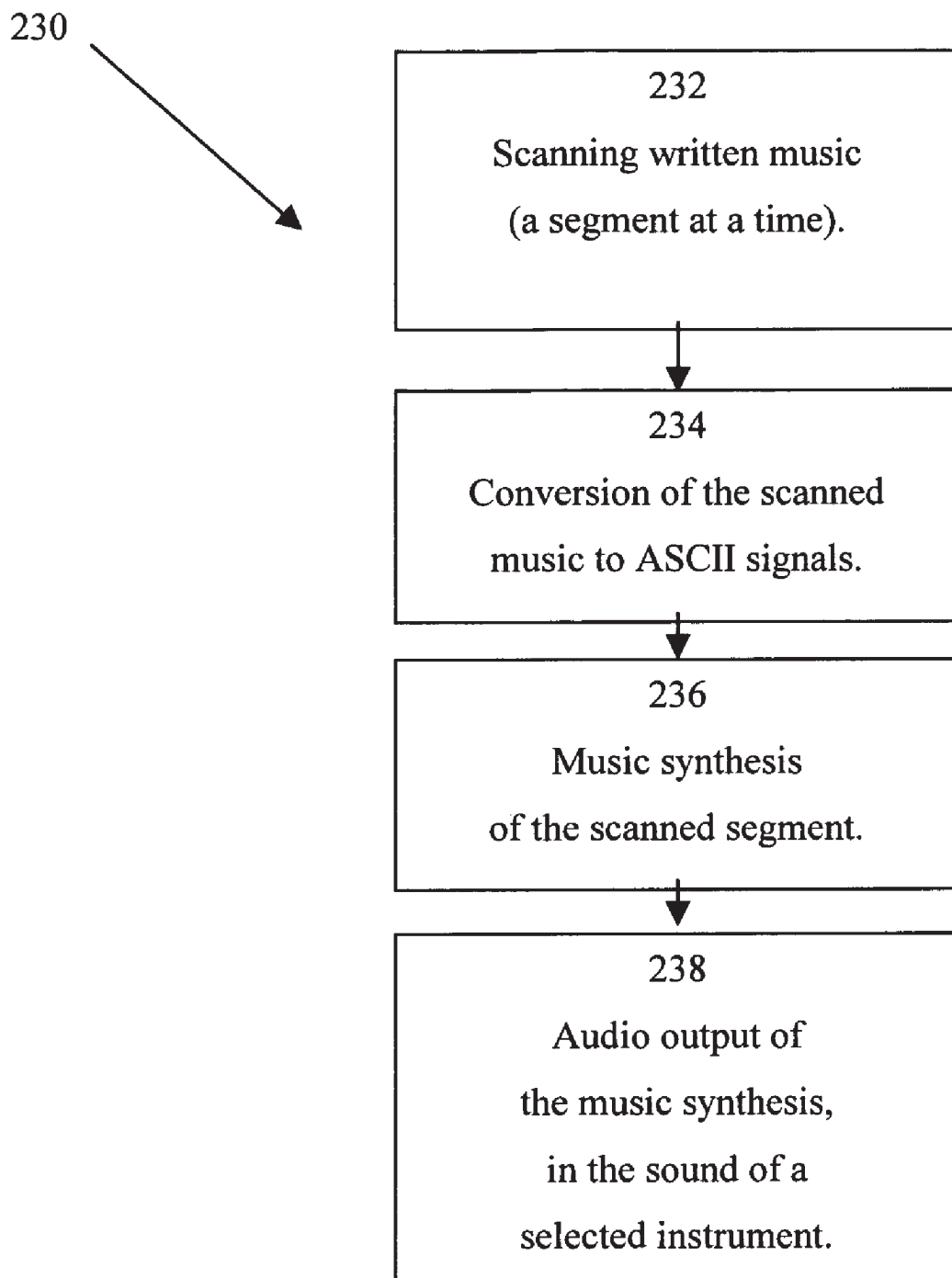


Fig. 9A

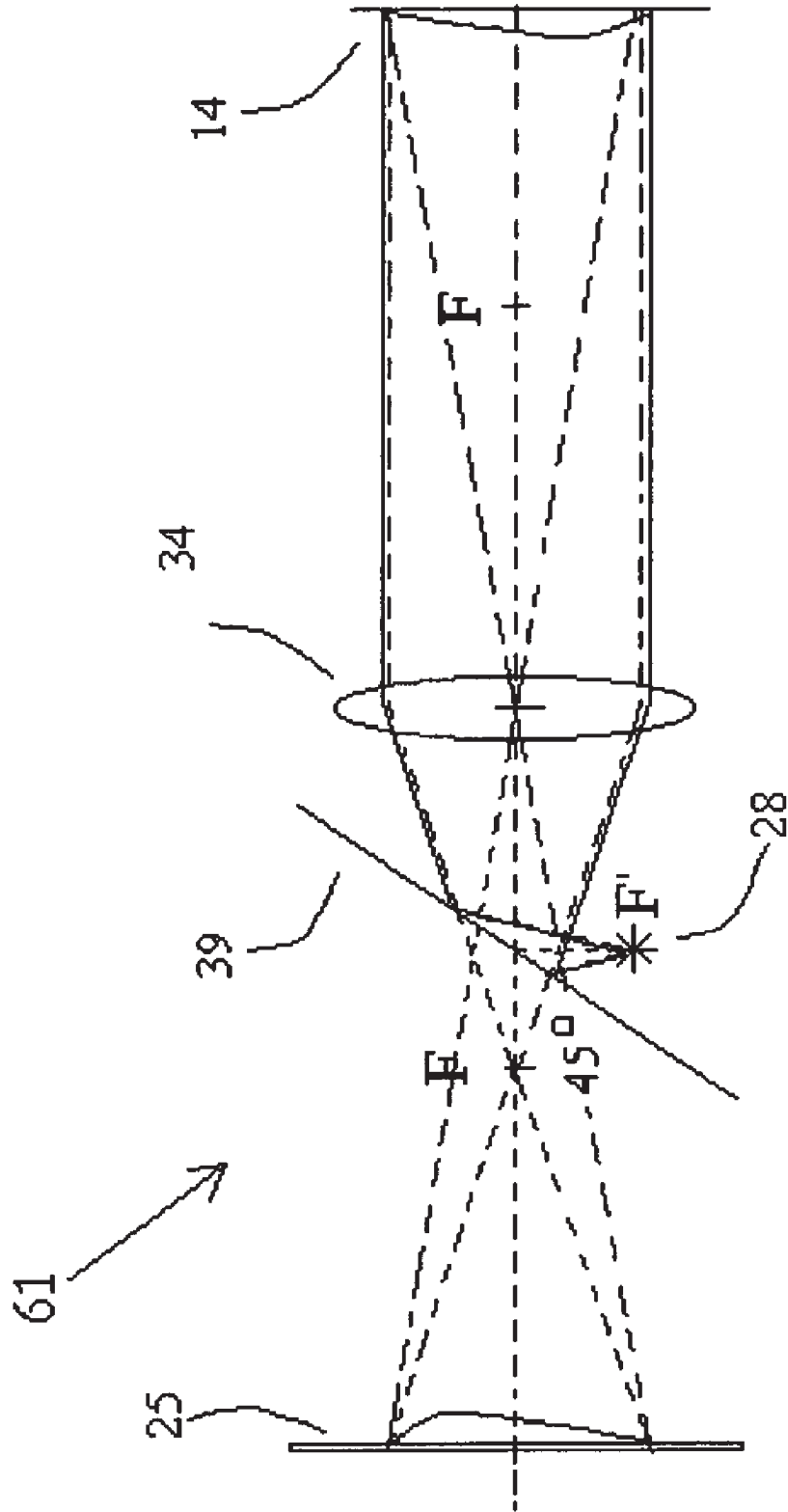


Fig. 9B

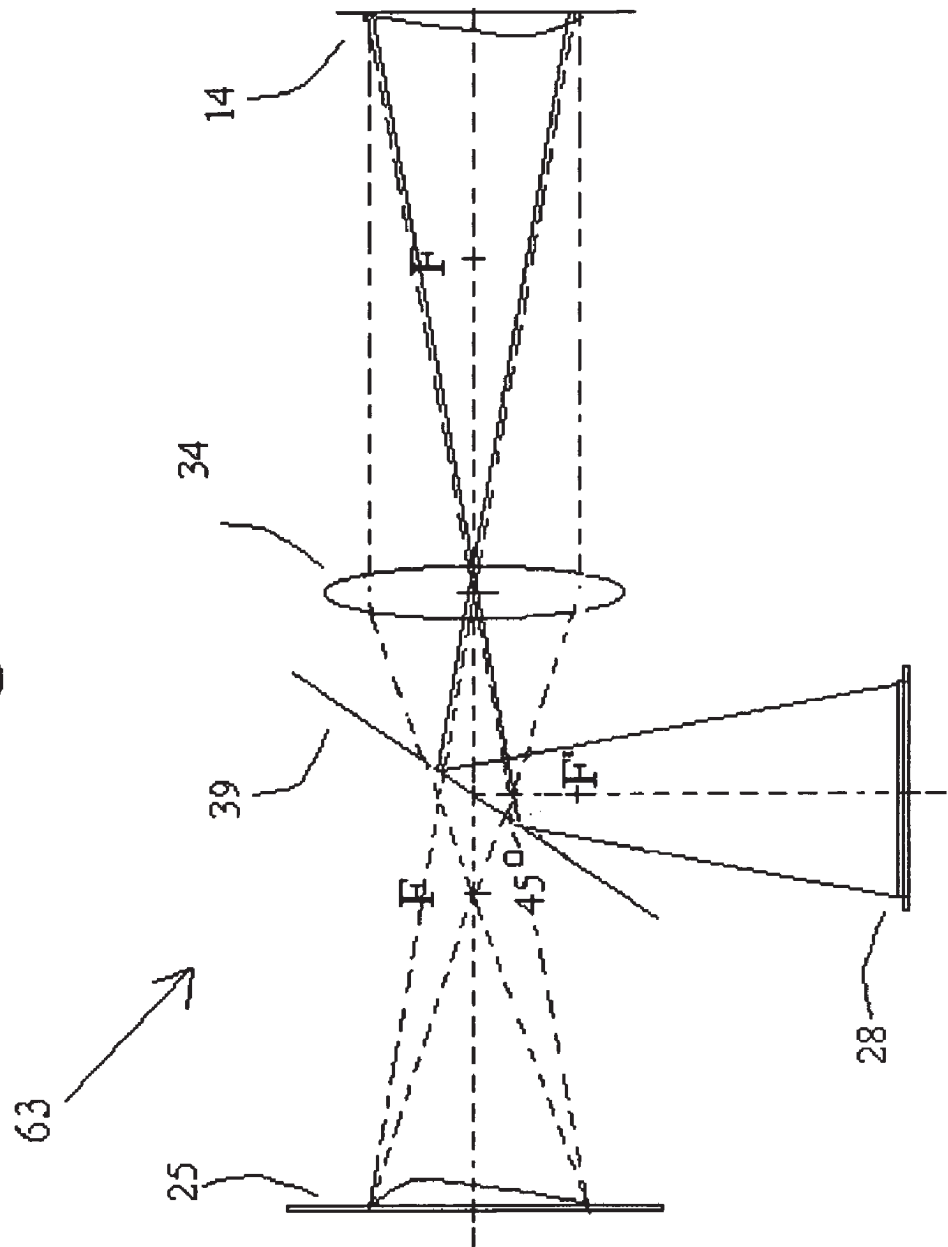
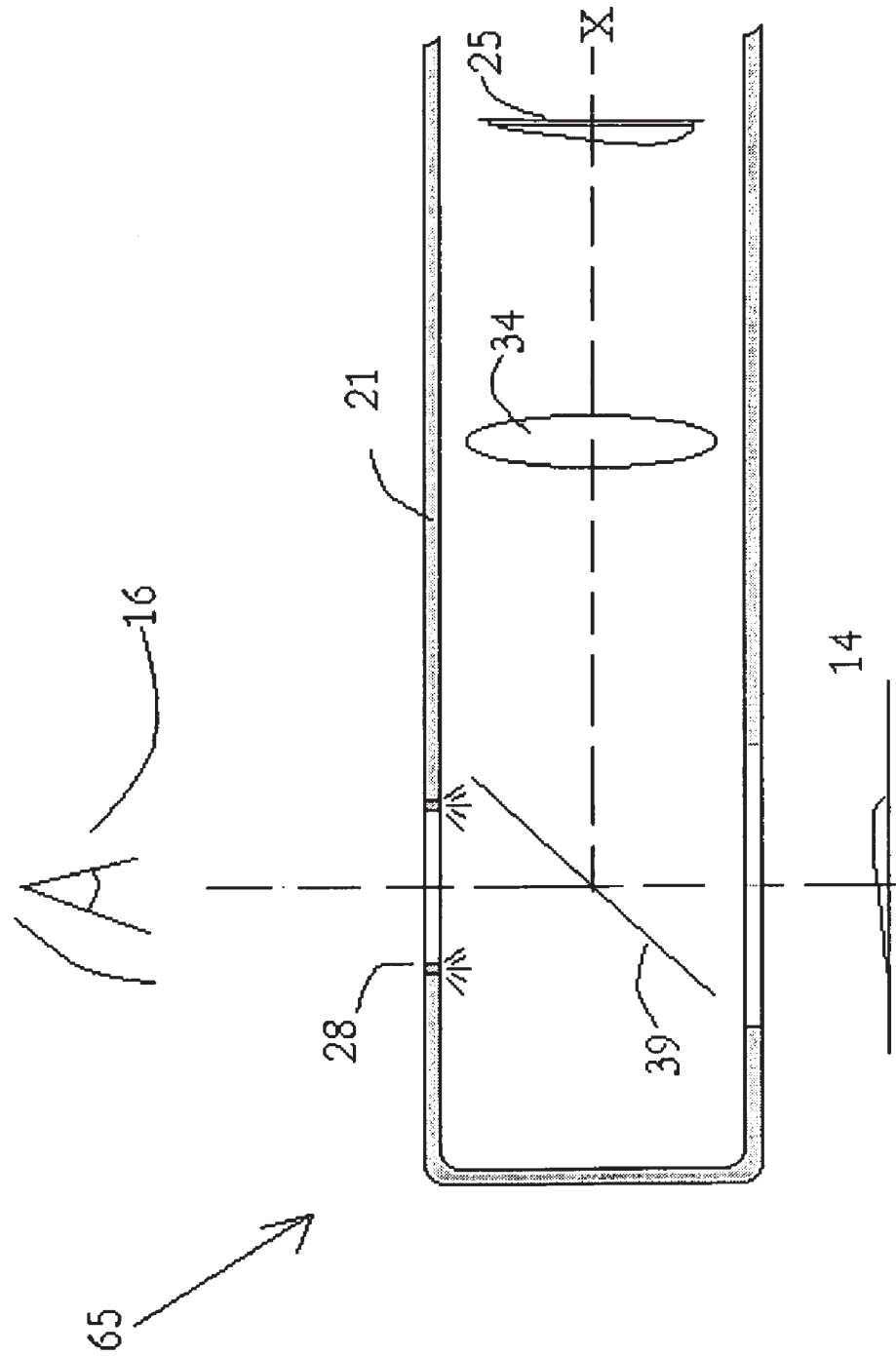


Fig. 9C



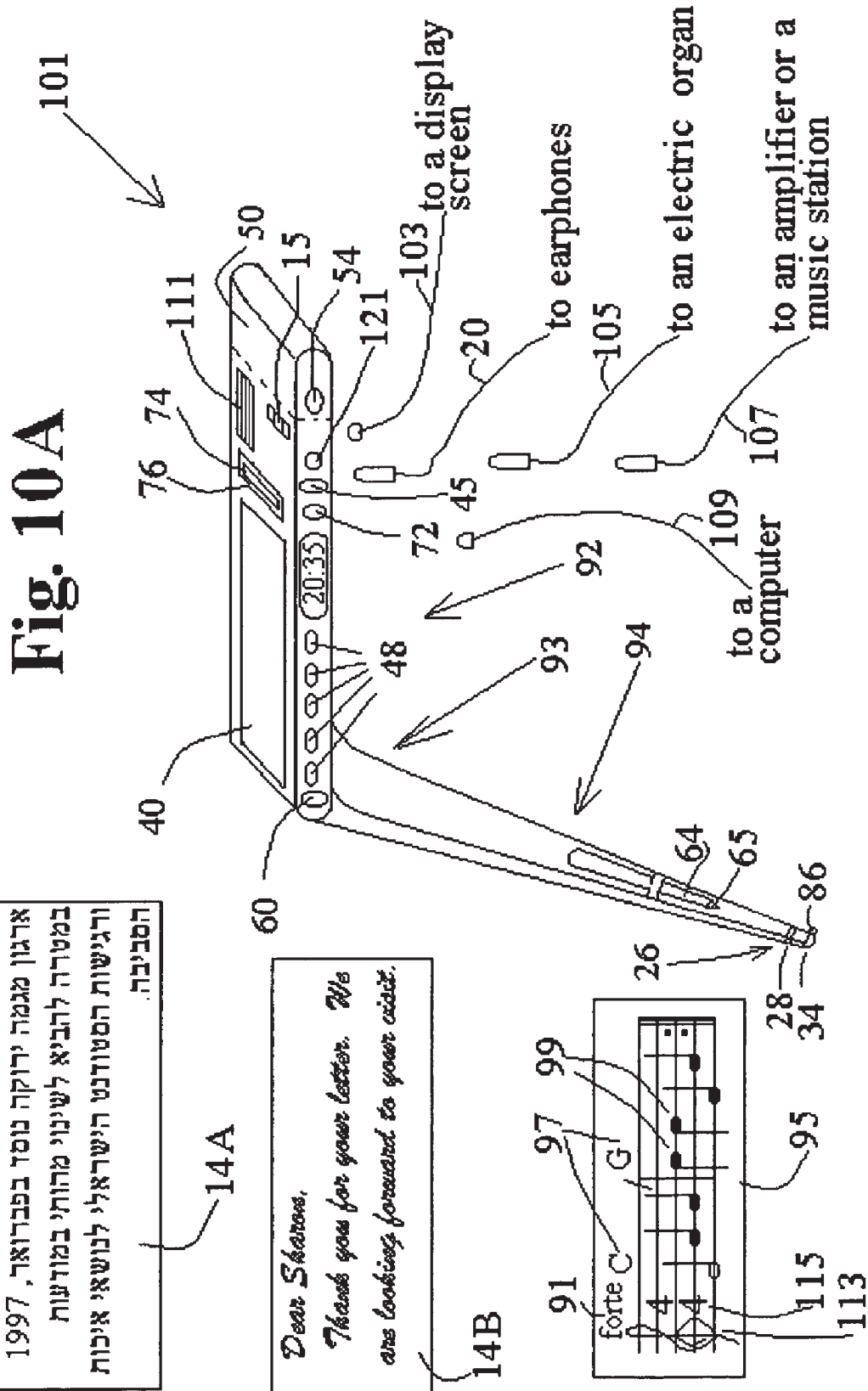


Fig. 10A

14A

ארגון מגמה ירוקה נוסד בפברואר, 1997 במטרה להביא לשינוי מהותי במודעות ורגישות הסטודנט הישראלי לנושאי איכות הסביבה.

14B

Dear Sharon,  
Thank you for your letter. We are looking forward to your visit.

91

forte C

4

97

99

26

28

34

95

115

113

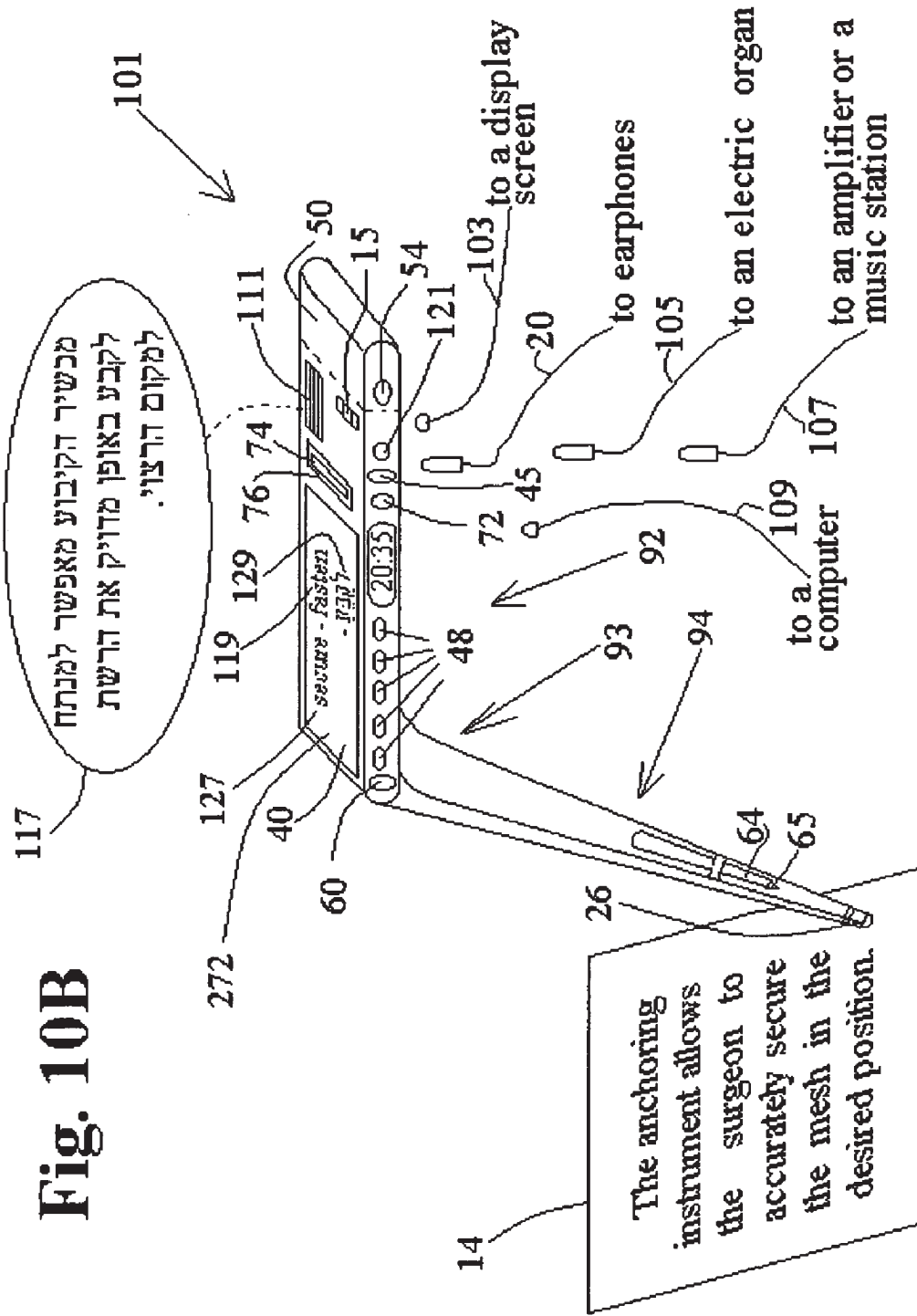
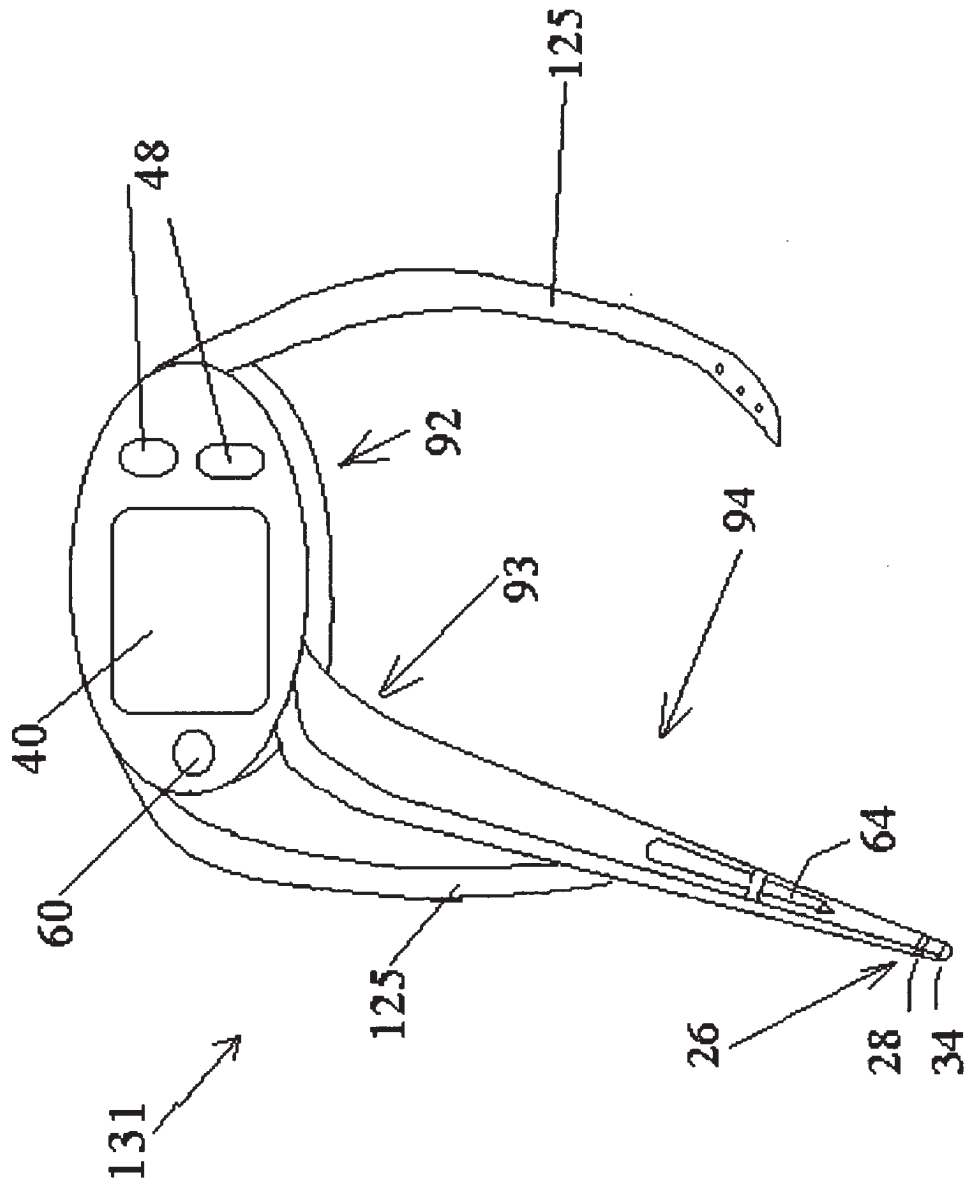


Fig. 10B



**Fig. 11**



**Fig. 12A**

<b>grammatical unit</b>	340 The local ski resort was opened to visitors on Saturday morning,				
<b>major functional component</b>	342 The local ski resort	344 was opened	346 to	348 visitors	350 on Saturday morning
<b>major component's function</b>	352 Subject: noun phrase	354 Predicate: verb phrase	356 preposition	358 object: noun tangible, human plural	360 adverbial phrase

**Fig. 12B**

<b>major functional component</b>	342 The local ski resort			
<b>basic functional component of dictionary-entry level</b>	362 The	364 local	366 ski	368 resort
<b>basic component's function</b>	372 article	374 adjective 1	376 adjective 2	378 noun
<b>basic component's attribute</b>	382 definite	384 adjective form	386 attributive noun form	388 tangible, inanimate, singular

**Fig. 12C**

390				
<i>General Adverbial-Phrase Template</i>				
<i>Preposition – Article - Adjective phrase(s) - Noun - Prepositional Phrase</i>				
<i>Preposition</i>	<i>Article</i>	<i>Adjective Phrase(s)</i>	<i>Noun</i>	<i>Prepositional Phrase</i>
By the river				
By	the	-	river	-
After dinner				
After	-	-	dinner	-
By his wisdom				
By	-	his	wisdom	-
With some luck				
With		some	luck	
In a minute				
In	a	-	minute	-
With a piercingly loud and screeching scream				
With	a	piercingly loud and screeching	scream	
By the rivers of Babylon				
By	the	-	rivers	of Babylon
In consequence of my smoking				
In consequence of			my	smoking

**Fig. 12D**

392 <i>General Imperative-Sentence Template</i> <i>Imperative verb – Preposition – Article - Adjective phrase(s) - Noun phrase</i>				
<i>Imperative Verb</i>	<i>Preposition</i>	<i>Article</i>	<i>Adjective Phrase(s)</i>	<i>Noun Phrase</i>
Fast forward a decade.				
Fast forward	-	a	-	decade
Return to sender.				
Return	to			sender
Write down your great ideas about your book.				
Write	down	-	your great	ideas about your book
Bring in your friend.				
Bring	in	-	your	friend
Go!				
Go				

**Fig. 13A**

<i>Function</i>	<i>Attributes, 1<sup>st</sup> Level</i>	<i>Attributes, 2<sup>nd</sup> Level</i>	<i>Attributes, 3<sup>rd</sup> Level</i>	<i>Attributes, 4<sup>th</sup> Level</i>
article				
	indefinite			
	definite			
noun				
	intangible			
		abstract		
			singular	
			plural	
		action		
			active	
				infinitive
				gerund
			passive	
				participle
	tangible			
		singular		
		plural		
			male	
			female	
				human
				animal
				vegetable
				object
				time

**Fig. 13B**

<i>Function</i>	<i>Attributes, 1<sup>st</sup> Level</i>	<i>Attributes, 2<sup>nd</sup> Level</i>	<i>Attributes, 3<sup>rd</sup> Level</i>	<i>Attributes, 4<sup>th</sup> Level</i>
noun (continued)				
	pronoun			
		I		
		you		
		he		
		she		
		it		
		we		
		they		
		me		
		you		
		him		
		her		
		us		
		them		
		myself		
		yourself		
		himself		
		herself		
		itself		
		ourselves		
		yourselves		
		themselves		
		mine		
		yours		
		his		
		hers		
		ours		
		yours		
		theirs		
		this		
		that		
		those		
		these		
		who		
		what		
		which		
		whose		

**Fig. 13C**

<i>Function</i>	<i>Attributes, 1<sup>st</sup> Level</i>	<i>Attributes, 2<sup>nd</sup> Level</i>	<i>Attributes, 3<sup>rd</sup> Level</i>	<i>Attributes, 4<sup>th</sup> Level</i>
adjective				
	adjective form			
	attributive noun form			
	gerund			
	participle			
	possessor			
adverb				
	adverbial form			
	question words			
		what		
		where		
		when		
		why		
		who		
		whose		
		whom		
		how		
		how come		
		which		
preposition	in			
	out			
	over			
	under			
	from			
	to			
	at			
	with			
	about			
	after			
	down			
	up			
	into			
	for			
	since			
	but			

**Fig. 13D**

<i>Function</i>	<i>Attributes, 1<sup>st</sup> Level</i>	<i>Attributes, 2<sup>nd</sup> Level</i>	<i>Attributes, 3<sup>rd</sup> Level</i>	<i>Attributes, 4<sup>th</sup> Level</i>
verb				
	intransitive			
	transitive			
		passive		
		active		
		imperative		
			present simple	
			present cont.	
			present perfect	
			past simple	
			past cont.	
			past perfect	
			future simple	
			future cont.	
			future going to	
			modal	may
				can
				could
				could have
				would
				would have
				should
				should have
				might
				might have
				ought to
				ought to have
				will have
				shall have



**Fig. 13E**

<i>Function</i>	<i>Attributes, 1<sup>st</sup> Level</i>	<i>Attributes, 2<sup>nd</sup> Level</i>	<i>Attributes, 3<sup>rd</sup> Level</i>	<i>Attributes, 4<sup>th</sup> Level</i>
conjunction				
	however			
	whatever			
	wherever			
	whenever			
	wherein			
	whereas			
	that			
	after			
	because			
	since			
	why			
	where			
	how			
	when			
	and			
	but			

**Fig. 14**

<b>430</b> <b>Exposure Basis</b>	
<b>440</b> <b>English</b>	<b>450</b> <b>Hebrew</b>
<p>The local ski resort was opened to visitors on Saturday morning, after some 70 centimeters of snow piled up in the area, in a snow storm that began Wednesday night. The resort management said that skiing will not be possible yet, because there was not enough snow on the mountain to enable safe skiing. Lower regions witnessed heavy rains and hail. As a result, rivers and streams filled up. The storm is expected to end on Saturday.</p>	<p>אתר התיירות המקומי נפתח למבקרים בשבת בבוקר, אחרי שכ-70 סמ' של שלג נערמו באזור, בעת סופת שלג שהחלה ביום רביעי בערב. הנהלת האתר מסרה כי עדין לא ניתן לגלוש מאחר ואין מספיק שלג על ההר לאפשר גלישה בטוחה. באזורים הנמוכים יותר ירדו גשמים עזים וברד. כתוצאה מכך, נהרות ונחלים התמלאו. הסופה אמורה לשכוח בשבת.</p>

**Fig. 15A**

Grammatical Units and Their Corresponding Functional Templates	
Grammatical Unit, 1 <sup>st</sup> language	442 The local ski resort
Functional Template, 1 <sup>st</sup> language	444 definite article, adjective 1 (adjective form), adjective 2 (attributive noun form), noun (tangible, object, singular).
Functional Template, 2 <sup>nd</sup> language	454 noun (tangible, object, singular, male), definite article, adjective 2 (possessive form), definite article, adjective 1 (adjective form).
Grammatical Unit, 2 <sup>nd</sup> language	452 אתר הסקי המקומי

**Fig. 15B**

Grammatical Units and Their Corresponding Functional Templates	
Grammatical Unit, 1 <sup>st</sup> language	442 The local ski resort
Functional Template, 1 <sup>st</sup> language	444 definite article, adjective 1 (adjective form), adjective 2 (noun form), noun (tangible, object, singular).
Functional Template, 3 <sup>rd</sup> language	464 definite article (female), noun (tangible, object, singular, female), adjective 2 [prepositional phrase (preposition, noun (male form, singular))], adjective 1 (adjective form, female).
Grammatical Unit, 3 <sup>rd</sup> language	462 La station de ski locale

**Fig. 15C**

Grammatical Units and Their Corresponding Functional Templates	
Grammatical Unit, 2 <sup>nd</sup> language	452 אתר הסקי המקומי
Functional Template, 2 <sup>nd</sup> language	454 noun (tangible, object, singular, male), definite article, adjective 2 (possessive form), definite article, adjective 1 (adjective form).
Functional Template, 3 <sup>rd</sup> language	464 definite article (female), noun (tangible, object, singular, female), adjective 2 [prepositional phrase (preposition, noun (male form, singular))], adjective 1 (adjective form, female).
Grammatical Unit, 3 <sup>rd</sup> language	462 La station de ski locale

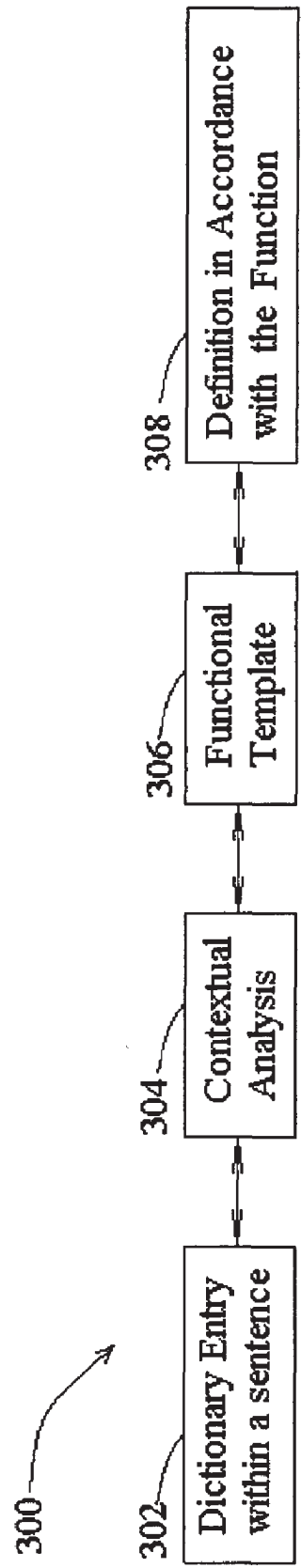


Fig. 16

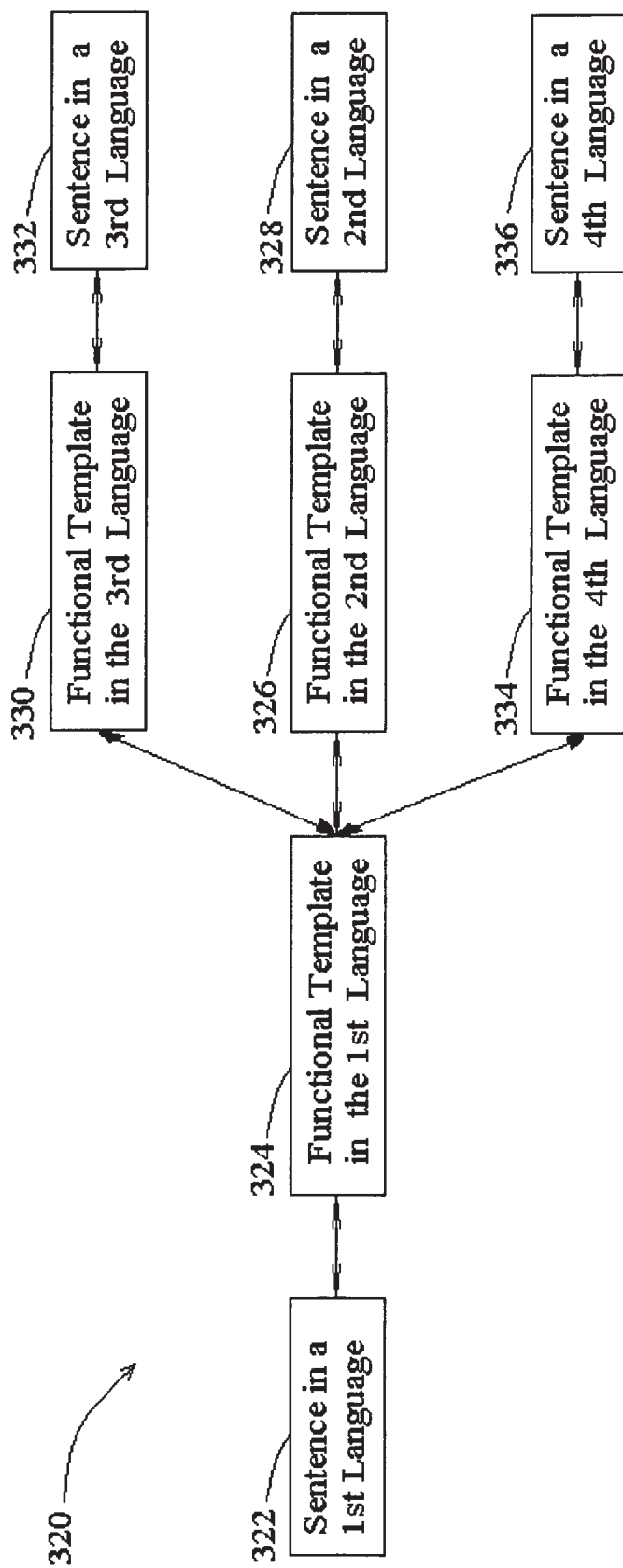


Fig. 17A

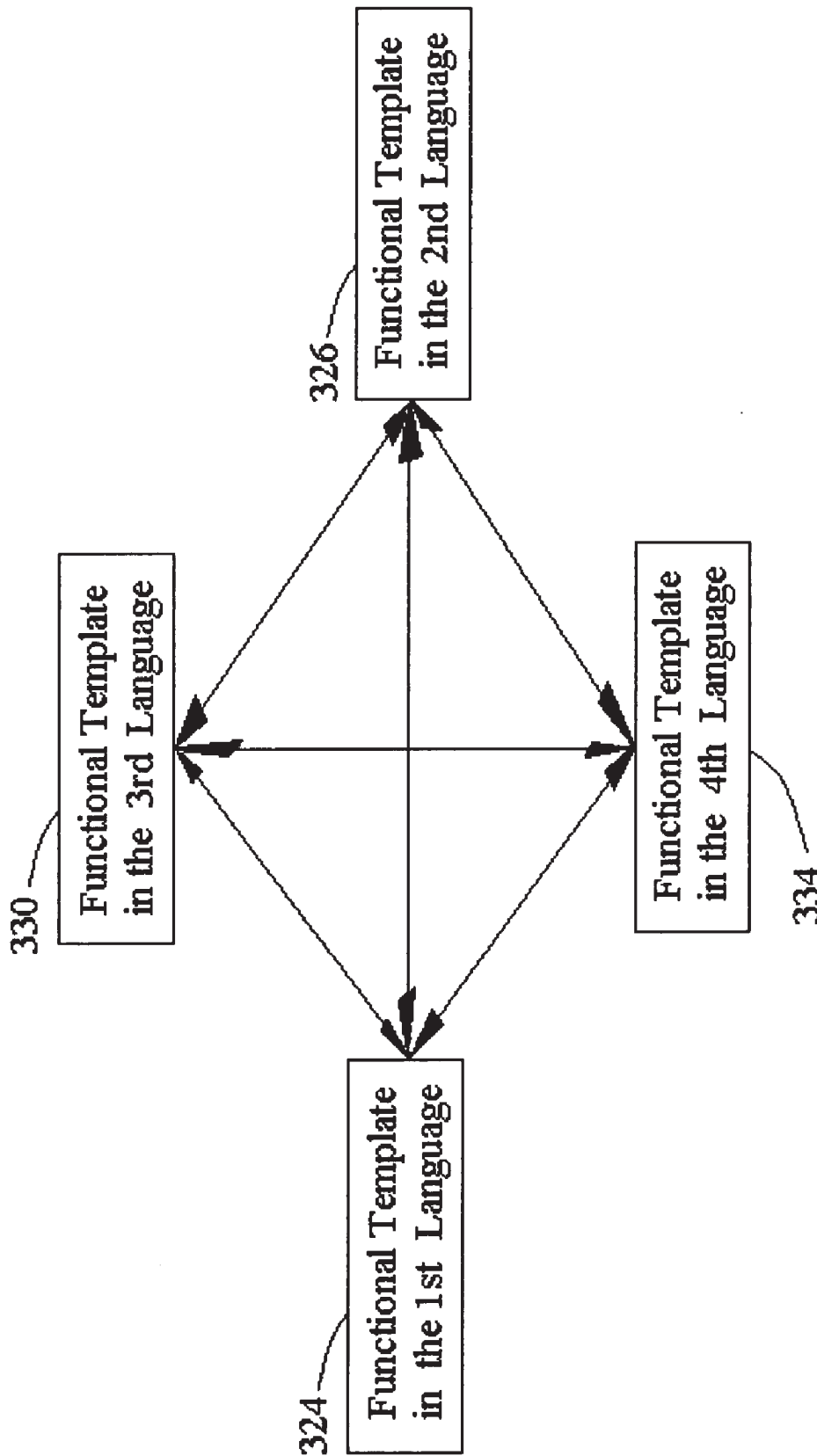
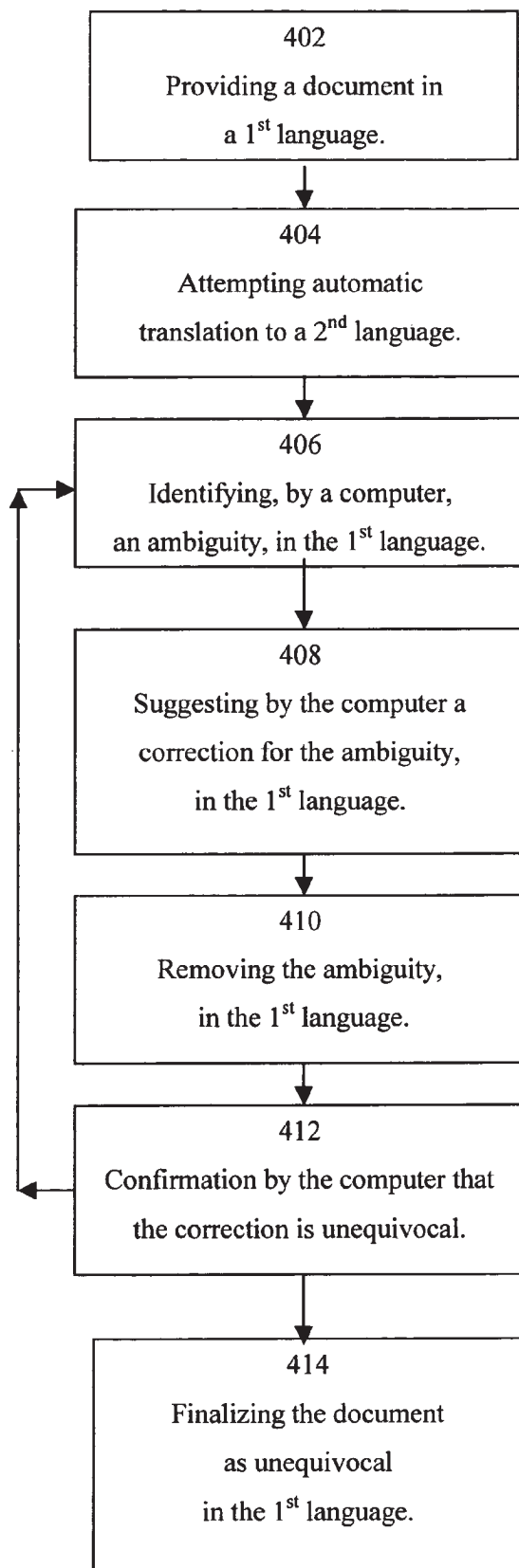


Fig. 17B



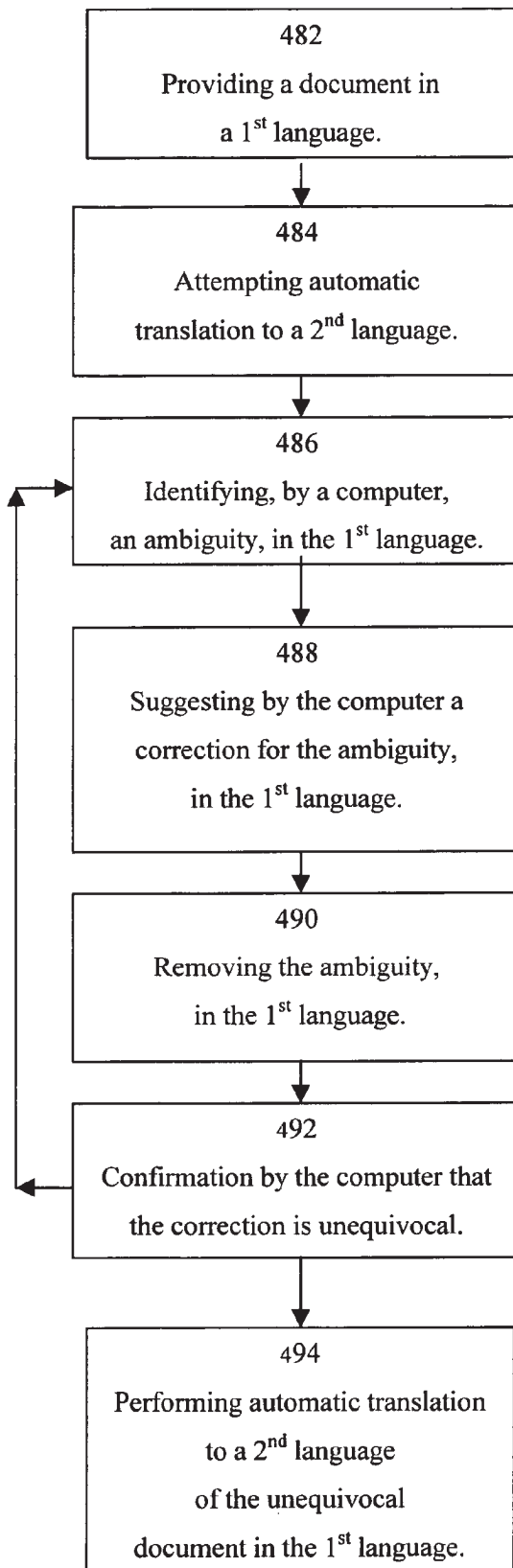
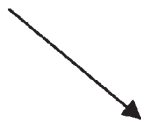
Fig. 18A

400



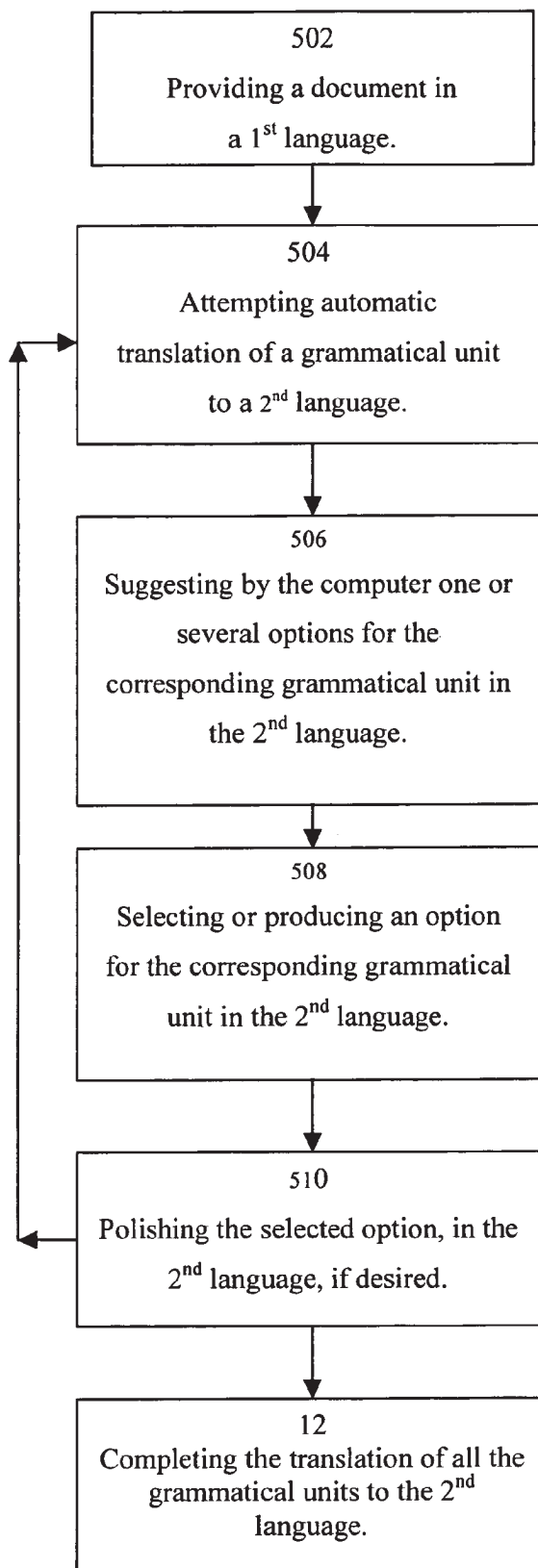
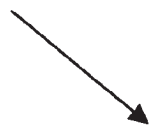
**Fig. 18B**

480



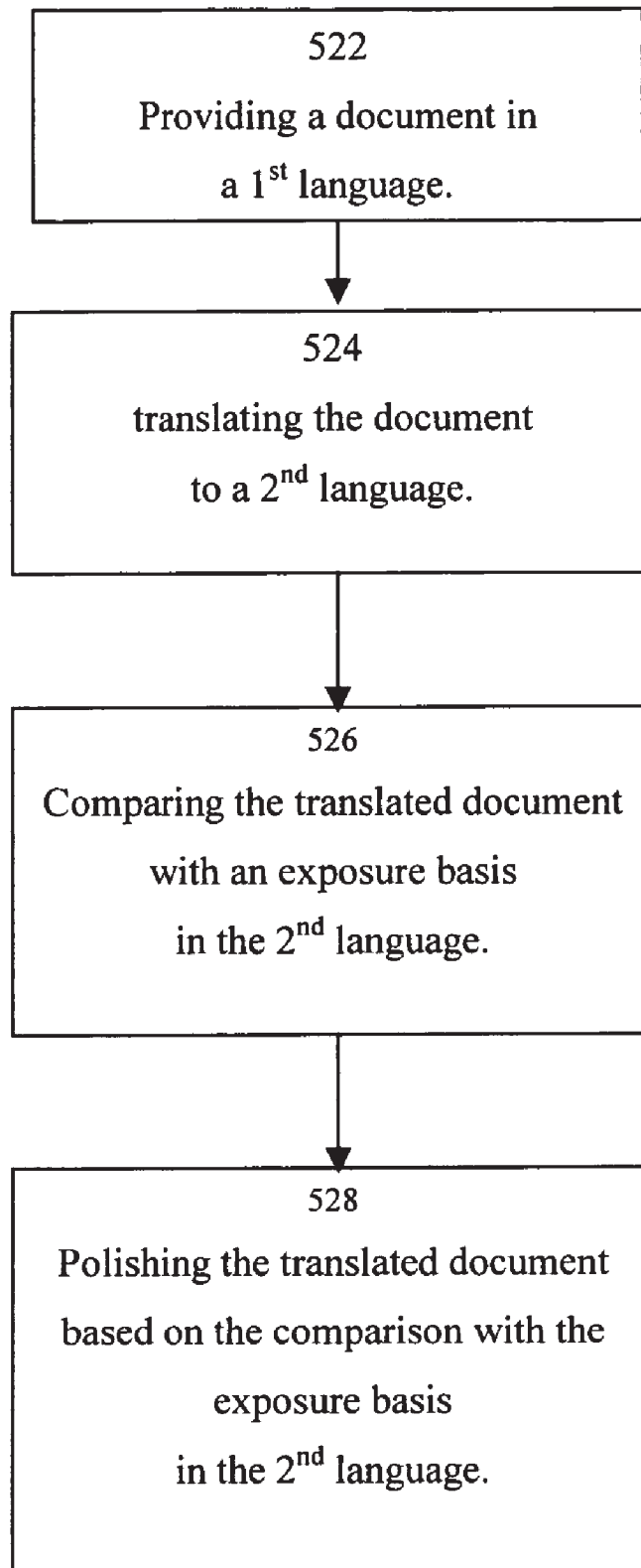
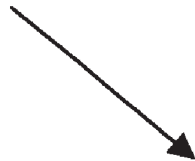
**Fig. 18C**

500



**Fig. 18D**

520



1

## IN-CONTEXT ANALYSIS AND AUTOMATIC TRANSLATION

### CROSS-REFERENCE TO RELATED APPLICATIONS

This is a Continuation-In-Part of U.S. patent application Ser. No. 09/946,391, filed on Sep. 6, 2001, now U.S. Pat. No. 6,704,699 which claims priority from U.S. Provisional Patent Application No. 60/229,794, filed on Sep. 5, 2000.

### FIELD OF THE INVENTION

The present invention relates to methods of presenting definitions in context and to methods of automatic translation, by computerized devices in general and by a stand-alone, hand-held, computerized scanning apparatus.

### BACKGROUND OF THE INVENTION

As instruments for acquiring a new language, dictionaries are sometimes ineffective. For example, dictionaries generally provide multiple meanings to words, and the meanings often depend on the functions the words play in a particular sentence. In Merriam-Webster OnLine Dictionary (<http://www.m-w.com>), the word "result" has two entries, as an intransitive verb and as a noun. The word "after" has six entries, as an adverb, a preposition, a conjunction, an adjective, a verbal auxiliary, and a noun. To look up a word, one must first select the appropriate function.

Yet, to a nonnative speaker of a language, determining the function of an unknown word in a specific sentence is not simple, since sentence structures in different languages differ. In English, for example, a sentence generally begins with adjectives to the subject, and some of these may be attributive nouns, operative as adjectives. Additionally, certain words may be used either as nouns or as verbs. In consequence, a list of words may be strung together, wherein each may function in two or more ways, so as to completely baffle a nonnative speaker. Consider for example:

"Stock market rates show improved performance."

Is "market" the subject, and "rates" the predicate? Is "rates" the subject, and "show" the predicate? Or is "show" the subject and "improved" the predicate?

Consider also,

"High health care costs result in poor health care availability."

Is "care" the subject, and "costs" the predicate? Or is "costs" the subject and "result" the predicate?

In fact, in some cases, determining the function of an unknown word in a sentence may be more difficult than inferring the meaning of a word, whose function in the sentence is known.

Native speakers are sensitive to subtle clues that generally make these sentences unequivocal. But nonnative speakers of a particular language may not be sufficiently familiar with or sufficiently tuned to these clues. Yet, it is on this point, crucial to understanding, that dictionaries offer no help.

Additionally, in most English dictionaries, definitions are provided for a word stem, when a verb, but not for its participle and gerund forms, which may function as adjectives. A certain linguistic skill is required for making the transformation to an adjective, a skill not always possessed by a nonnative speaker of English, whose native tongue does not include such transformations. Furthermore, since ordinary English dictionaries generally do not treat the participle

2

and gerund forms as adjectives, they do not explain the distinction between them, when functioning as adjectives.

The use of attributive nouns is another cause for puzzlement. Attributive nouns relate to two or more nouns, juxtaposed, with no connecting words. Is "a science fair" a fair of things scientific, or a science of organizing fairs? Neither the American Heritage® Dictionary of the English Language: Fourth Edition, 2000, <http://www.bartleby.com/61/>, nor the Merriam-Webster OnLine Dictionary provide an answer.

Even for a specific function, dictionaries often provide a plurality of meanings. For example, the word "order" as a transitive verb, has six meanings, according to the American Heritage® Dictionary. Yet, in a particular context, many of these meanings are irrelevant, and their inclusion makes the use of the dictionary cumbersome, even frustrating.

An instrument, which has the power of a dictionary, but which is specifically adapted for language acquisition, is desired.

### SURVEY OF PRIOR ART

Optical scanners are known. They convert objects such as pictures, barcodes, or portions of text to machine-readable data signals. Typically, the data signals are read by a user's computer to reproduce an image of the scanned object on a display device, such as a CRT, a display screen or a printer.

A hand-held optical scanner is manipulated by hand across the object that is being scanned. The hand-held scanner may be connected directly to the user's computer by a data cable, and may transfer image data to the computer as the data are collected. Alternatively, the hand-scanner may be a stand-alone unit and may include a data storage component for storing the image data. The data may be downloaded to a separate computer after the scanning operation is complete.

A hand-held optical scanner generally includes an illumination system, for illuminating the region to be scanned, an optical system, for collecting and focusing light reflected by the illuminated, scanned region, a photosensitive detector, for detecting the light collected and focused thereon by the optical system, an analog amplifier, for amplifying the signals produced by the photosensitive detector, and an analog-to-digital converter, for converting the amplified signals to digitized machine-readable data signals. The illumination system may be, for example, a fluorescent or incandescent lamp or an array of light emitting diodes (LEDs). The optical system may include a lens or a lens-and-mirror assembly.

The photosensitive detector is generally a Charge-Coupled Device (CCD). A CCD includes an array of photosensitive cells, or pixels, each pixel collecting an electrical charge responsive to the light that falls upon it. Thus, a CCD may be used to detect light and dark spots of a scanned object. The charge from each pixel is converted to an analog voltage by an analog amplifier, and the analog voltage is digitized by an Analog-to-Digital Converter (ADC). The digitized signals are the machine-readable data signals, which can be stored or processed by the user on a computer or a similar device.

Sometimes, a Contact Image Sensor (CIS) is used in place of the CCD. In a CIS scanner, the array of photosensitive cells is arranged in close proximity to the object to be scanned, so as to catch the reflected light directly; an optical system is not necessary.

U.S. Pat. No. 5,996,895 to Heiman, et al, incorporated herein by reference, describes a scanning system with adjustable light output and/or scanning angle.

U.S. Pat. No. 6,033,086 to Bohn, incorporated herein by reference, describes a compact illumination system for a hand-held scanner.

U.S. Pat. No. 5,841,121 to Koenck, incorporated herein by reference, describes a hand-held optical scanner, having automatic focus control, for operation over a range of distances.

U.S. Pat. No. 5,019,699 to Koenck, incorporated herein by reference, describes a hand-held optical scanner, which includes a lens system having circular symmetry. The lens system focuses the full width of the object onto an array of photosensitive cells, with a single flash of a ring-type xenon flash tube, which surrounds the lens system and is symmetrically arranged relative to the optical axis. In this way, the object can be scanned at any angle relative to the array of photosensitive cells, and the scanned image, stored in digital form, can be electronically rotated to a desired orientation, before it is decoded.

U.S. Pat. No. 5,834,749 to Durbin, incorporated herein by reference, describes a hand-held scanner for reading images at oblique angles, in order for the scanning unit not to interfere with the user's view of the scanned image. The distortion to an obliquely scanned image, arising from the oblique scanning, can be corrected by any of several correction techniques, as follows:

1. a ratio of vertical to horizontal line densities of the array of photosensitive cells can be chosen to compensate for the vertical foreshortening of the scanned image;
  2. the array of photosensitive cells can be oriented at an oblique angle with respect to the optical axis, to compensate for the distortion inherent in the oblique scanning;
  3. a lens system can be configured to provide varying degrees of magnification along its surface; and
- as taught by U.S. Pat. No. 5,019,699, to Koenck, described hereinabove, processing techniques can electronically re-orient the scanned image after storing it in the scanner's memory.

Hand-held, stand-alone, optical scanners that produce audio output are known. U.S. Pat. No. 5,945,656 to Lemelson, et al, incorporated herein by reference, describes a pen-like stand-alone scanner for transducing coded data into pre-coded pieces of speech or music. Generally, a scanning guide is attached to a book, arranged for guiding the pen-like scanner vertically along an edge of the book, which contains coded information. Aided by the guide, children may scan the coded data and produce the sounds associated with them.

U.S. Pat. No. 5,767,494 to Matsueda, et al., incorporated herein by reference, describes a system for reproducing multimedia information, recorded with an optically readable code. The code is a dot-code format described in U.S. Ser. No. 08/407,018 (PCT Publication No. WO 94/08314), and includes two-dimensional patterns that convey multimedia information, for example, audio information like speech and music, image information obtained from a camera or a video device, and digital code data obtained from the user's computer. The system uses paper as a basic information-storage medium for the optically readable code, and includes a preferably pen-like scanner, arranged to read the code. The system may reproduce the original multimedia information by appropriate hardware such as a display screen, a printer, or a speaker, and includes a speech synthesizer. In some embodiments, the pen-like scanner is a stand-alone unit, and may include earphones.

The IRISPen of Image Recognition Integrated Systems Inc., of Rue Du Bosquest 10, 1348 Louvain-la-Neuve, Belgium, is a pen-like scanner that allows the user to scan text, bar codes and handwritten numbers into any Windows

or Mac application. The IRISPen is hooked up to any desktop or portable computer without any additional interface boards. The IRISPen is not stand-alone apparatus.

The IRISPen Executive™ integrates text-to-speech technology from Lernout & Hauspie, in six languages (English, French, German, Dutch, Spanish and Italian). It provides natural audio feedback of all recognized words and numbers as it scans the information. The purpose of the text-to-speech technology is to reduce the need of the user to keep his eyes on the computer screen to verify recognition. The IRISPen Translator is further arranged to automatically translate text between English and German. Output may be in the form of written text in the translated language, displayed on a computer screen or printed. Alternatively, the output may be an audio output, in the translated language.

The IRISPen Executive™, the IRISPen Translator, and other IRISPen products are not stand-alone apparatus. Rather, they are arranged to operate with a computer, such as a desktop PC or a notebook computer, into which the IRISPen software has been installed. The output language is the language that has been installed to the computer, and cannot be changed during a scanning operation. An audio output may be provided only in the original language or only in the translated language. Furthermore, the automatic translation language is not intrinsic to the IRISPen Translator. Rather, it has to be installed on the computer that supports the IRISPen. Neither are the speaker or earphones for audio output intrinsic to the IRISPen. Instead, the computer speakers are used for the audio output. Therefore, the IRISPen is not a single product but a package of several products, which are sold together and are arranged to operate together.

Text-to-speech (TTS) syntheses, additional to the technology of Lemout & Hauspie of the IRISPen Executive™, are known. Bell Labs and Edinburgh University have developed a text-to-speech synthesis based on a Spoken Text Markup Language (STML) standard. STML later became SABLE. Sun Microsystems, Inc., in partnership with other speech-technology companies, has worked to define the specifications for a Java Speech API and a Java Speech Markup Language (JSML), incorporating many of the aspects of SABLE. JSML has been accepted by W3C (the organization responsible for WWW standards) as a standard. Bell Labs Lucent Technologies now offer a text-to-speech synthesis, which provides choices between voices of a man, a woman or a child and a speech rate that is fast, normal or slow. The University of Edinburgh has developed a generally multi-lingual system known as The Festival Speech Synthesis System, available in English (British and American), Spanish and Welsh. Additionally, Digital offers the DECTalk™ Speech Synthesizer which converts ASCII text to natural-sounding speech output. IBM offers the V5.1 speech synthesizer. Apple offers "English Text-to-Speech" software with recent versions of the MacOS. The University of York has produced YorkTalk, Oxford University offers an all-prosodic speech synthesizer entitled IPOX. Telcordia Technologies (formerly Bellcore) have developed the ORATOR and an improved version, the ORATOR II. Entropic Research Laboratory, Inc. offers TrueTalk 1.0, a software-only text-to-speech system based on a major research effort at AT&T Bell Laboratories. AT&T has developed Next-Generation TTS to convert machine-readable English text into audible speech. The Speech Technology Unit at BT has produced, and is continuing to develop, a sophisticated text-to-speech system called Laureate. Eurovocs is still another commercially available, text-to-speech product. BORIS is a high-quality, diphone-based text-to-speech converter for Spanish, developed by Universidad Politecnica de