

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC
Petitioner,

v.

SOUNDCLEAR TECHNOLOGIES LLC
Patent Owner.

IPR2025-00344
U.S. Patent No. 9,070,374

**DECLARATION OF MR. STUART LIPOFF,
UNDER 37 C.F.R. § 1.68 IN SUPPORT OF PETITION FOR
INTER PARTES REVIEW**

TABLE OF CONTENTS

I.	Introduction.....	9
II.	Qualifications and Professional Experience.....	11
III.	Relevant Legal Standards	18
	A. Level Of Ordinary Skill.....	20
	B. Prior Art.....	20
	C. Anticipation	21
	D. Obviousness.....	21
IV.	'374 Patent	26
	A. Overview	26
	B. Prosecution History	28
V.	Technology Background.....	29
	A. Speech Processing	29
	B. User Feedback	31
VI.	Level of Ordinary Skill in the Art	34
VII.	Claim Construction.....	34
	A. “simplex communication apparatus”	35
VIII.	Claims 1-15 are Unpatentable	36
	A. Ground 1: IBM Renders Claims 1-3, 5, 9, 10, and 13 Obvious.....	37
	1. IBM Summary	37

2.	Claim 1	39
a)	[1pre] A condition notification method for notifying a used condition of a simplex communication apparatus by using a light-emitting device attached to the simplex communication apparatus, the method comprising:	39
b)	[1a] a communication-mode determination step of determining whether a communication mode of the simplex communication apparatus is a transmission mode or a standby mode;	40
c)	[1b] a sound pick-up state determination step of determining a sound pick-up state of a sound carried by a speech signal to be transmitted if the communication mode is the transmission mode;	42
d)	[1c] a control step of controlling the light-emitting device so that the light-emitting device is turned off, turned on or repeatedly turned on and off based on determination results of the communication-mode determination step and the sound pick-up state determination step;	44
e)	[1d] a speech-quality evaluation step of evaluating speech quality of the speech signal to be transmitted, ...	46
f)	[1e] wherein, in the sound pick-up state determination step, the sound pick-up state of the sound carried by the speech signal to be transmitted is determined if the communication mode is the transmission mode based on the speech quality of the speech signal evaluated by the speech-quality evaluation step; and	50

- g) [1f] wherein, in the sound pick-up state determination step, the sound pick-up state of the sound carried by the speech signal to be transmitted, if the communication mode is the transmission mode, is determined based on a determination result of the speech-segment determination step and an evaluation result of the speech-quality evaluation step. 53
- h) Claim 2: The condition notification method according to claim 1, wherein, in the control step, the light-emitting device is turned on if the communication mode is the transmission mode, whereas the light-emitting device is repeatedly turned on and off based on the sound pick-up state determined in the sound pick-up state determination step if the communication mode is the transmission mode.....55
- 3. Claim 3: The condition notification method according to claim 1, wherein, in the speech-quality evaluation step, the speech quality of the speech signal to be transmitted is evaluated based on a volume level of a sound to be picked up. 57
- 4. Claim 5.....58
 - a) [5a] The condition notification method according to claim 1, wherein in the speech-segment determination step it is determined whether or not the speech signal to be transmitted is a speech segment if the communication mode is the transmission mode, wherein, ...58
 - b) [5b] in the speech-quality evaluation step, the speech quality of the speech signal to be transmitted is evaluated while the speech signal to be transmitted is being determined as the speech segment. 59

5.	Independent Claim 9	60
a)	[9pre] A communication apparatus comprising:	60
a)	[9a] a first pick-up unit configured to pick up a voice sound;	60
b)	[9b] a transmitter unit configured to transmit the voice sound picked up by the first pick-up unit to outside as a first speech signal;	61
c)	[9c] a communication-mode switching unit configured to switch a communication mode between a standby mode in which the transmitter unit does not transmit the speech signal and a transmission mode in which the transmitter unit transmits the speech signal;	62
d)	[9d] a sound pick-up state determination unit configured to determine a pick-up state of the voice sound picked up by the first pick-up unit;	64
e)	[9e] a light emission device configured to emit light;	64
f)	[9f] a control unit configured to control the light-emitting device so that the light-emitting device is turned off, turned on or repeatedly turned on and off based on the communication mode switched by the communication-mode switching unit, and the pickup state of the voice sound picked up by the first pick-up unit determined by the sound pick-up state determination unit; and.....	65
g)	[9g] a speech-quality evaluation unit configured to evaluate speech quality of the first speech signal to be transmitted by the transmitter unit, ...	66

- h) [9h] wherein the sound pick-up state determination unit determines the sound pick-up state of the voice sound picked up by the first sound pick-up unit based on the speech quality of the speech signal evaluated by the speech-quality evaluation unit; and 68
 - i) [9i] a speech-segment determination unit configured to determine whether or not the first speech signal to be transmitted by the transmitter unit is a speech segment, ...68
 - j) [9i] wherein, the sound pick-up state determination unit determines the sound pick-up state of the sound to be transmitted as the first speech signal based on a determination result at the speech-segment determination unit and an evaluation result at the speech-quality evaluation unit. 69
6. Claim 10: The communication apparatus according to claim 9, wherein the light-emitting device is turned on if the communication mode is switched into the transmission mode, whereas the light-emitting device is repeatedly turned on and off based on the pick-up state determined by the sound pick-up state determination unit if the communication mode is switched into the transmission mode.....69
7. Claim 13.....70
- a) [13a] The communication apparatus according to claim 9, wherein the control unit controls the light-emitting device so that the light-emitting device is turned off if the communication mode is the standby mode, ...70
 - b) [13b] the light-emitting device is turned on if the communication mode is the transmission mode and if the speech signal to be transmitted is not the speech segment, the

	light-emitting device is turned on if the communication mode is the transmission mod, if the speech signal to be transmitted is the speech segment, and if the speech quality of the speech signal to be transmitted is evaluated as bad, and the light-emitting device is repeatedly turned on and off if the communication mode is the transmission mode, if the speech signal to be transmitted is the speech segment, and if the speech quality of the speech signal to be transmitted is evaluated as good.....	71
B.	Ground 2: IBM and Kale Render Obvious Claims 4, 6-8, 11, 12, and 15	72
1.	Kale Summary	72
2.	Reasons to Combine IBM and Kale (Ground 2)	74
3.	Claims 4 and 11	75
a)	Claim 4.....	75
b)	Claim 11.....	78
4.	Claim 7.....	80
5.	Claims 6 and 12	85
a)	Claim 6.....	85
b)	Claim 12.....	89
6.	Claims 8 and 15	90
a)	Claim 8.....	90
b)	Claim 15.....	95
C.	Ground 3: IBM, Kale, and Li Render Claim 14 Obvious	96
1.	Overview of Li.....	96

2.	Reasons to Combine	98
3.	Claim 14: The communication apparatus according to claim 11, wherein the first and second sound pick-up units are a main microphone and a sub-microphone, respectively, and the main microphone and the sub-microphone are arranged at a front face and a rearface of the communication apparatus, respectively.	100
IX.	Conclusion	102

I, Stuart Lipoff, do hereby declare as follows:

I. Introduction

[0001] I make this declaration based upon my own personal knowledge and, if called upon to testify, would testify competently to the matters stated herein.

[0002] I have been asked by Google LLC (“Google”) to provide my expert opinion connection with the *Inter Partes* Review of U.S. Patent No. 9,070,374 (“the ’374 patent”) to Kanevsky et al. concerning whether the ’374 patent is unpatentable over certain prior art. This declaration is a statement of my opinions on issues relating to the patentability of claims 1-15 (“the Challenged Claims”) of the ’374 patent. As I explain more fully below, it is my opinion that all of the Challenged Claims would have been obvious to a person of ordinary skill in the art (“POSITA”) at the time of the alleged invention.

[0003] I am being compensated for my work in this matter at my standard hourly rate. I am also being reimbursed for reasonable and customary expenses associated with my work and testimony in this investigation. My compensation is not contingent on the outcome of this matter or the specifics of my testimony.

[0004] I have reviewed and considered the following documents in connection with my analysis of the '374 patent:

Exhibit	Description
Ex. 1001	U.S. Patent No. 9,070,374 to Konishi ("the '374 patent")
Ex. 1002	File History of the '374 Patent
Ex. 1005	U.S. Patent Application Publication No. US 2008/0167868 A1 to Kanevsky et al., assigned to International Business Machines ("IBM")
Ex. 1006	U.S. Patent Appl. Pub. No. 2012/0027219 to Kale et al. ("Kale")
Ex. 1007	U.S. Patent Application Publication No. 2011/0026730 to Li et al. ("Li")
Ex. 1008	International Telecommunication Union (ITU) Terms and Definitions ("simplex operation")
Ex. 1009	J. D. Hoyt and H. Wechsler, "Detection of human speech in structured noise," <i>Proceedings of ICASSP '94. IEEE International Conference on Acoustics, Speech and Signal Processing</i> , Adelaide, SA, Australia (1994).
Ex. 1010	Joon-Hyuk Chang, Nam Soo Kim and S. K. Mitra, "Voice activity detection based on multiple statistical models," in <i>IEEE Transactions on Signal Processing</i> , vol. 54, no. 6, pp. 1965-1976, June 2006.
Ex. 1011	VU Metter, Newton's Telecom Dictionary (16th ed. 2000).
Ex. 1012	"The Authoritative Dictionary of IEEE Standards Terms, Seventh Edition," in <i>IEEE Std 100-2000</i> , vol., no., pp. 1-1362, 11 Dec. 2000,
Ex. 1013	Alex Artaud, <i>Warm up your sound with this tiny well-featured monster</i> , <i>Electronic Musician Magazine</i> , vol. 17, no. 8, pp. 144-146, August 2001.

[0005] In forming the opinions expressed below, I have considered:

- a) the documents listed above;
- b) the relevant legal standards, including the standard for obviousness, and any additional authoritative documents as cited in the body of this declaration; and
- c) my own knowledge and experience based upon my work in the field of software systems as described below, as well as the following materials.

[0006] I reserve the right to supplement and amend any of my opinions in this declaration based on documents, testimony, and other information that becomes available to me after the date of this declaration.

[0007] Unless otherwise noted, all emphasis in any quoted material has been added.

II. Qualifications and Professional Experience

[0008] I am currently president of IP Action Partners Inc., a consulting practice that serves the telecommunications, information technology, media, electronics, and e-business industries.

[0009] I earned a Bachelor of Science degree in Electrical Engineering in 1968 and a second Bachelor of Science degree in Engineering Physics in 1969, both from Lehigh University. I earned a

Master of Science degree in Electrical Engineering from Northeastern University in 1974, and then a Master of Business Administration degree from Suffolk University in 1983.

[0010] I hold a Federal Communications Commission (“FCC”) General Radiotelephone License. I also hold a Certificate in Data Processing from the Institute for the Certification of Computing Professionals (“ICCP”), which is supported by the Association for Computing Machinery (“ACM”).

[0011] I am also a registered professional engineer (PE) in the Commonwealth of Massachusetts and in the State of Nevada.

[0012] I am a fellow of the Institute of Electrical and Electronics Engineers (“IEEE”) Consumer Electronics, Communications, Computer, Circuits, and Vehicular Technology Groups. I have been a member of the IEEE Consumer Electronics Society National Board of Governors (formerly known as the Administrative Committee) since 1981, and I was Boston Chapter Chairman of the IEEE Vehicular Technology Society from 1974 to 1976. I served as the 1996-1997 President of the IEEE Consumer Electronics Society, and from 1999 to 2018 I served as Chairman of the Society’s Technical Activities and Standards Committee and as Vice President of Publications for the Society. From 2018 to 2023 I served as Vice President of Standards and Industry Activities for the Society and

currently serve on the Board of Governors as The Historian for The Society. I have also served as an Ibuka Award committee member for the IEEE's Award in the field of consumer electronics.

[0013] I have prepared and presented numerous papers at the IEEE and at other professional meetings. For example, in fall 2000, I served as general program chair for IEEE's Vehicular Technology Conference on advanced wireless communication technology. I have organized sessions at The International Conference on Consumer Electronics, and I was the 1984 program chairman. I conducted an eight-week IEEE-sponsored short course on Fiber Optics System Design. I received IEEE's Centennial Medal in 1984, and I received IEEE's Millennium Medal in 2000.

[0014] As Vice President and Standards Group Chairman for the Association of Computer Users ("ACU") from 1980 to 1983, I served as the ACU representative to the ANSI X3 Standards Group. From 1976 to 1978, I served as Chairman of the task group on user rule compliance for the FCC's Citizens Advisory Committee on Citizen's Band Radio.

[0015] Over the last 25 years, I have been a member of the Society of Cable Television Engineers, the Association for Computing Machinery, and The Society of Motion Picture and Television Engineers. From 2001 to 2004, I served as a member of the USA advisory board to the National

Science Museum of Israel. In 1998, I presented a short course on international product development strategies as a faculty member for Technion Institute of Management in Israel. From 2001 to 2003, I served as a member of the board or directors of The Massachusetts Future Problem Solving Program.

[0016] I am a named inventor on seven United States patents and have several publications on data communications in publications, including Electronics Design, Microwaves, EDN, the Proceedings of the Frequency Control Symposium, Optical Spectra, and IEEE publications.

[0017] During my professional career dating from 1969 to the present, I have been heavily engaged in the study, analysis, evaluation, design, and implementation of products and technology associated with consumer electronics and electronic appliances. A particular focus of my professional activities has been improving the man-machine interface including voice, speech, and speaker recognition for man-machine interactions. I also have extensive experience in studying foundation technologies and the applications supporting speech signal processing.

[0018] For approximately three years, from 1969 to 1972, I served as Project Engineer for Motorola's Communications Division, where I had project design responsibilities for paging and wireless communication

products. Projects that I worked on while employed at Motorola included work on paging systems that included digital voice storage, voice compression, and voice synthesis. I also worked on projects that interfaced wireless data communications terminals to public safety computer systems for mobile data retrieval and data entry.

[0019] For approximately four years, from 1972 to 1976, I served as Section Manager for Bell & Howell Communications Company, where I also had project design responsibilities for paging and wireless communication products. The projects I supported included covert audio intelligence systems that recognized speech and activated digital voice compression recording systems. I also led projects for voice-based radio paging systems that recorded speech input, processed the speech to remove silence, processed the speech to digitally compress the speech, and store and forward the speech upon demand from DTMF or computer keyboard retrieval from the servers.

[0020] For 25 years from 1976 to 2001, I worked for Arthur D. Little, Inc. (ADL), where I became the Vice President and Director of Communications, Information Technology, and Electronics (CIE) and served in that role for 10 years, from 1991 to 2001. At ADL, I was responsible for the firm's global CIE practice in laboratory-based contract

engineering, product development, and technology-based consulting. I was also involved in multiple pioneering efforts to identify and explore customer-to-business and business-to-business electronic commerce and transactions information processing opportunities (e-commerce). These projects involved technology assessment and analysis as well as developing architectures and systems to support multiple applications, and typically involved an information retrieval component.

[0021] While at ADL, I worked on several projects involving the combination of voice interfaces (including speech recognition and voice audio output) and information retrieval. For example, over the course of three years in the early-1990s, I worked on a project for Bolt Beranek and Newman (BB&N), where I evaluated and benchmarked technology for a voice input/output application that allowed end users (*e.g.*, travel agents) to use speech inputs to interact with airline reservation databases to retrieve information about travel reservation options, where the results were returned to the user in an audible message. This system included a natural language front-end speech-interface module with speech recognition that used pre-defined recognition grammars to convert the end user's speech into structured commands supported by an airline reservation system. As another example, over the course of three years in the mid-1990s, I worked on a

project for Texas Instruments that applied a speech–recognition interface for a variety of applications that retrieved information from database servers.

[0022] Other projects that I worked on at various points in my 25 years at ADL and afterwards involving speech recognition technologies included the following.

[0023] Over the course of three years in the early 1990s, I worked on a voice–interface project developing spoken digit telephone number recognition and voiceprint matching for Sprint’s long distance alternative access telephone services.

[0024] Over the course of a year in the late 1980s, I worked on a voice interface project evaluating the processing power needed to perform various voice recognition applications by Rockwell Semiconductor’s signal processing technology.

[0025] Over the course of 15 years starting in the early 1980’s, I worked on a project for the United States Postal Service (USPS), where we developed a real-time automated postal teller system that served as an interface between end-users and the USPS’s information systems. This system included voice prompts for the vision impaired.

[0026] I also have extensive experience in public and private network wired and wireless voice telecommunications while employed by

Motorola, Bell & Howell, and Arthur D Little, and while self-employed. In the course of these telecommunications projects ranging from 1969 to the present, I have encountered a number of applications where audio input and voice are used to activate devices, for example for the purpose of saving battery power by entering into low power, so-called “sleep” modes. These projects have involved the design of cellular telecommunications systems that implement industry standard means of entering lower power modes in the absence of voice.

[0027] I understand my curriculum vitae will become provided as Exhibit 1004.

III. Relevant Legal Standards

[0028] I am an engineer and not a lawyer. My understanding of the legal standards to apply in reaching the conclusions in this declaration is based on discussions with counsel for Petitioner, my experience applying similar standards in other patent-related matters, and my reading of the documents submitted in this proceeding. I have applied these legal standards in preparing this declaration.

[0029] I have been informed that there are two ways in which prior art may render a patent claim unpatentable. First, I have been informed that the prior art can “anticipate” a claim. Second, I have been informed that the

prior art can render a claim “obvious” to a POSITA. I understand that a claim is patentable if it was not anticipated and would not have been rendered obvious by the prior art at the effective filing date of the patent.

[0030] I have been informed that a dependent claim is a patent claim that refers back to another patent claim. I have been informed that a dependent claim includes all of the limitations of the claim to which it refers plus its own limitation(s).

[0031] I have been asked to provide my opinions as to whether the cited prior art discloses or renders obvious claims 1-15 of the '374 Patent from the perspective of a POSITA at the time of the earliest claimed priority date of February 20, 2012, as described in more detail below.

[0032] I have been informed that in IPR proceedings, such as this one, the party challenging the patent bears the burden of proving unpatentability by a preponderance of the evidence. I understand that a preponderance of the evidence means “more likely than not.”

[0033] For purposes of this declaration, I have been asked to provide my opinions on issues regarding unpatentability. I have been informed of the following legal standards, which I have applied in forming my opinions.

A. Level Of Ordinary Skill

[0034] I have been informed that a POSITA is determined by considering several factors, including the (i) type of problems encountered in the art; (ii) prior art solutions to those problems; (iii) rapidity with which innovations are made; (iv) sophistication of the technology; and (v) educational level of active workers in the field.

[0035] I have been instructed to assume that a POSITA is not a specific real individual, but rather a hypothetical individual having the qualities reflected by the factors discussed above. A POSITA is assumed to be person of ordinary creativity familiar with the prior art as of the priority date of the patent at issue.

B. Prior Art

[0036] I have been advised and understand that the information used to evaluate whether an invention was new and not obvious when made is generally referred to as “prior art.” I understand that in an IPR proceeding, prior art includes patents and printed publications that existed before the earliest claimed priority date or the earliest filing date of the patent (which I have been informed is also called the “effective filing date”). I have been informed and understand that a patent or published patent application is prior art if it was filed before the earliest filing date of the claimed invention and

that a printed publication is prior art if it was publicly available before the earliest filing date.

C. Anticipation

[0037] I have been informed that under 35 U.S.C. § 102, a patent claim is unpatentable for anticipation if the claimed subject matter was patented or described in a printed publication before the effective filing date of the claimed invention. I have been informed that this is referred to as unpatentability by anticipation. I have been informed that a patent claim is anticipated under § 102 if a single prior art reference discloses all the limitations of the claimed invention. I understand that limitations may be expressed or inherent such that the limitation is essential to the prior art.

D. Obviousness

[0038] I have been informed that for obviousness under 35 U.S.C. § 103, a patent claim is unpatentable if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious to a POSITA to which said subject matter pertains at the time the invention was made. I have been informed that this is referred to as unpatentability by obviousness.

[0039] I have been informed that an obviousness analysis includes the following considerations:

- a) Determining the scope and content of the prior art;
- b) Ascertaining the differences between the prior art and the claims at issue;
- c) Resolving the level of ordinary skill in the pertinent art; and
- d) Considering evidence of secondary indicia of nonobviousness (if available).

[0040] I have been informed that the relevant time for considering whether a claim would have been obvious to a POSITA is the time of invention. For my obviousness analysis, counsel for Petitioner instructed me to assume that the date of invention for the challenged claims is February 20, 2012. My opinions would not change if I assumed another, e.g., later, date of invention.

[0041] I have been informed that a reference may be modified or combined with other references or with a POSITA's own knowledge if the person would have found the modification or combination obvious. I have also been informed that a POSITA is presumed to know all the relevant prior art, and the obviousness analysis may take into account the inferences and creative steps that a POSITA would employ.

[0042] I have been informed that an obviousness determination must be made from the perspective of a POSITA. I have also been informed

that there is no requirement that the prior art contain an express suggestion to combine known elements to achieve the claimed invention, and that a suggestion to combine known elements to achieve the claimed invention may come from the prior art as a whole or individually. Also, the obviousness analysis may rely on the inferences and creative steps a POSITA would employ, as filtered through his or her knowledge as of the priority date. But I understand that obviousness grounds cannot be sustained by mere conclusory statements and must include some articulated reasoning and rationale to support a legal conclusion of obviousness.

[0043] In determining whether a prior art reference could have been combined with another prior art reference or other information known to a POSITA, I have been informed that the following principles may be considered:

- a) A combination of familiar elements according to known methods is likely to be obvious if it yields predictable results;
- b) The substitution of one known element for another is likely to be obvious if it yields predictable results;
- c) The use of a known technique to improve similar items or methods in the same way is likely to be obvious if it yields predictable results;

- d) The application of a known technique to a prior art reference that is ready for improvement to yield predictable results;
- e) Any need or problem known in the field and addressed by the reference can provide a reason for combining the elements in the manner claimed;
- f) A person of ordinary skill often will be able to fit the teachings of multiple references together like a puzzle; and
- g) The proper analysis of obviousness requires a determination of whether a POSITA would have a “reasonable expectation of success”—but not “absolute predictability” of success—in achieving the claimed invention by combining prior art references.

[0044] I have been informed that, when a work is available in one field, design alternatives and other market forces can prompt variations of it, either in the same field or in another. I have been informed that if a POSITA could have implemented a predictable variation and would have seen the benefit of doing so, that variation is likely to have been obvious. I have been informed that, in many fields, such as the mechanical or electrical arts, market demand—not scientific literature—may drive design trends. I have been informed that, when there was a design need or market pressure and

there are a finite number of predictable solutions, a POSITA would have had a good reason to pursue those known options.

[0045] I have been informed that the law permits the application of “common sense” in examining whether a claimed invention would have been obvious to a POSITA. For example, I have been informed that combining familiar elements according to known methods and in a predictable way may suggest obviousness when such a combination would yield nothing more than predictable results. I understand, however, that a claim is not obvious merely because every claim element is disclosed in the prior art. A party asserting obviousness must provide a specific motivation to combine or modify the references as recited in the claims and explain why one skilled in the art would have reasonably expected to succeed in doing so.

[0046] I have been informed that there is no rigid rule that a reference or combination of references must contain a “teaching, suggestion, or motivation” to combine references. But I also understand that the “teaching, suggestion, or motivation” test can be a useful guide in establishing a rationale for combining elements of the prior art. I have been informed that this test poses the question as to whether there is an express or implied teaching, suggestion, or motivation to combine prior art elements in a way that results in the claimed invention, and that it helps to counter the

use of hindsight, which is impermissible. Likewise, if a prior art reference “teaches away” from a potential prior art combination, then a motivation to combine may not exist.

[0047] I am not aware of any evidence of secondary considerations, such as unexpected results, industry skepticism, long-felt unresolved need, commercial success, praise by others, or copying that would alter my opinions set forth below.

[0048] I have been informed that, in an obviousness analysis, prior art must be analogous art to the patent being considered. I have been informed that a prior art reference is considered to be analogous, or in the same field of art, if the reference is either (1) in the same field of endeavor as the challenged patent, regardless of the problems the challenged patent and the prior art address, or (2) reasonably pertinent to the particular problem being solved by the challenged patent.

IV. '374 Patent

A. Overview

[0049] The '374 patent describes an audio input device that provides noise-cancellation capabilities and visual feedback to user on the device's processing status for captured speech. '374 patent, 24:6-18, 36:56-37:8, 37:33-47. The '374 patent explains the input device's feedback reflects

a “sound pick-up state determination” of received speech and indicates whether the microphone system is being utilized correctly. ’374 patent, 24:6-18, 36:56-37:8, 37:33-47.

[0050] The ’374 patent describes the “sound pick-up state determination” as the combined outcomes of traditional speech-segment and speech-quality determinations made for speech received at its microphone system according to communication mode. ’374 patent, 19:29-35. To switch between communication modes, the ’374 patent relies on button presses by the user on conventional Push-to-Talk (PTT) microphones. ’374 patent, 4:37-5:5, 6:6-23, 34:14-27, 39:24-33, Fig. 1 (104), Fig. 4 (207). Otherwise, to determine the sound pick-up state of captured speech, the ’374 patent explains:

[F]or sound pick-up state evaluation, the sound pick-up state evaluation unit 1003 determines whether a signal input from the sound pick-up unit 1001 [i.e., a microphone] is a speech segment. Moreover, for sound pick-up state evaluation, the sound pick-up state evaluation unit 1003 evaluates the speech quality of a speech segment to be transmitted from the transmitter unit 1002. The results of speech-segment determination and speech-quality evaluation are output to the light-emission control unit 1005
’374 patent, 19:29-35.

[0051] The ’374 patent goes on to explain that its associated “light-

emitting device 1210 [is] for visual notification,” including to “notify a user of speech quality” and “on correct holding of the main microphone.” ’374 patent, 26:2, 26: 51-52, 33:17-19. And at Figure 13, the ’374 patent provides a corresponding table listing exemplary LED states based on the communication mode and sound pick-up state determination (i.e., the speech segment and speech quality determinations):

	COMMUNICATION MODE	SOUND PICK-UP STATE DETERMINATION		LED STATE
		SPEECH SEGMENT DETERMINATION	SPEECH QUALITY DETERMINATION	
No1	STANDBY MODE (NO TRANSMISSION INSTRUCTION SIGNAL INPUT)	NO DETERMINATION	NO DETERMINATION	TURN OFF
No2	TRANSMISSION MODE (INPUT TRANSMISSION INSTRUCTION SIGNAL)	NON-SPEECH SEGMENT	NO DETERMINATION	TURN ON
No3		SPEECH SEGMENT	BAD SPEECH QUALITY	TURN ON
No4			GOOD SPEECH QUALITY	TURN ON AND OFF

’374 patent, Fig. 13

[0052] The ’374 patent recognizes that these on-off sequences for a given LED state simply represent one way to convey the underlying “sound pick-up state.” ’374 patent, 39:50 (“the present invention is not limited thereto.”).

B. Prosecution History

[0053] In response to a first-action obviousness rejection, the applicant amended independent claims 1 and 11 to incorporate certain

subject matter indicated allowable aside from their dependence on a rejected independent claim. Ex. 1002, 26-33. In particular, the applicant incorporated corresponding dependent claims 5 and 14, respectively, requiring that a “speech-segment determination” inform the “sound pick-up state determination.” Ex. 1002, 26-33. The applicant also incorporated intervening dependent claims 3 and 13, which required a speech-quality evaluation to also inform the sound pick-up state determination. Ex. 1002, 26-33. The Examiner’s subsequent Notice of Allowance, however, still focused on the incorporated subject matter of dependent claims 5 and 14—i.e., inclusion of a speech-segment determination as part of the sound pick-up state. Ex. 1002, 12-13.

V. Technology Background

A. Speech Processing

[0054] Processing audio signals for speech determinations goes back decades in the industry, as does providing feedback regarding those determinations. Speech detection—also referred to as voice activity detection (VAD)—evaluates audio inputs having a mix of speech and noise to determine whether there is a quality component of the audio input identifiable as speech. VAD technology has long been the subject of many academic publications, recognizing the inherent importance of accurately

distinguishing human speech segments from noisy signals. A 1994 paper coauthored by the FBI, titled “Detection of Human Speech in Structured Noise,” directly addressed the concept:

This paper describes research to develop an efficient system that provides a binary decision as to the presence of speech in a short (one to three second) time sample of an acoustic signal. A method which is efficient and reliably detects human speech in the presence of structured noise (such as wind, music, traffic sounds, etc.) is described.

Ex., 1009, Abstract. As explained, “[t]he problem area of interest in this research is in providing a binary yes / no decision that human speech is present in a time segment of an acoustic signal of one to three seconds in duration.” *Id.*, Introduction. Underlying the motivations even then, the paper recognized that “simple commands and queries via speech rather than a keyboard would be preferable” for future communication devices. *Id.*

[0055] Similarly, evaluating speech quality for enhancement is a related, longstanding interest within the industry. In another IEEE article titled “Voice Activity Detection Based on Multiple Statistical Models,” the authors addresses VAD challenges on this point. Ex. 1010. As the paper explains:

Voice activity detection (VAD) refers to the classical problem of distinguishing active speech from nonspeech and has applications for a variety of speech communication systems, such as speech coding, speech recognition, noisy speech enhancement, hands-free conference, and echo cancellation

Ex. 1010, Introduction. The paper goes on to describe techniques for improving speech quality through “noisy speech enhancement” based on selection of a one or more different statistical models for better distinguishing speech from noise, especially at speech transition regions. *Id.*

B. User Feedback

[0056] Providing user feedback about received audio signals has also been a longstanding practice in the industry. A common form of earlier visual indicators of audio quality include the Volume Unit (VU) Meter, which Newton’s Telecom Dictionary defines in 2000 as:

VU is the unit of measurement for electrical speech power in communications work VUs are measured in decibels above 1 milliwatt. The measuring device is called a VU meter, which is an abbreviation of a volume-unit meter, a type of meter used to indicate average audio amplitude.

Ex. 1011, 3. Similarly, The Authoritative Dictionary of IEEE Standards Terms provides a complementary definition:

[S]tandard volume indicator (volume measurements of electrical speech and program waves) A device for the indication of volume, and having the characteristic described in IEEE Std 152-1953w. Note: A standard volume indicator consists of at least two parts: 1) A meter; and 2) an attenuator (adjustable loss) or pad (fixed loss).

Ex. 1012.

[0057] Initially, VU meters were implemented as a mechanical meter with a moving needle, while lighted displays later became a preferred means of providing visual feedback for speech and/or speech quality determinations. Some devices used both during the transition. For example, in August 2001, Electronic Musician published an article describing a mechanical VU meter produced by ART that included a supporting LED indicator regarding indicating quality. Ex. 1013.



ART Tube MP Studio Tube Microphone Preamp with VU Meter and Output Protection Limiter, https://www.bhphotovideo.com/c/product/1713863-REG/art_tubempst_tube_mp_studio_mic.html/overview (last visited December 31, 2024); *see also* Ex. 1013, 2.

[0058] The LED indicators provide additional visual feedback regarding speech quality. In particular:

Directly below the VU meter is an LED that glows green when power is applied and remains green with lower-level signals. If OPL is not in use, the LED turns red to indicate clipping. However, the occasional flashes of red didn't result in unpleasant distortion. The LED works as a good reference for calibrations

(after your ears, of course); a constantly lit red LED means the signal level needs to be reduced. With OPL engaged, the LED glows red when the limiter circuitry senses that the signal is about to clip and thus begins reducing gain. *Id.*

[0059] Thus, a variety of feedback systems for audio processing devices existed as of 2001. But as I explain below with respect to the petition grounds, the industry continued to develop more advanced feedback systems as supporting technology allowed one of skill in the art to engage more options for providing the processing status of the underlying audio device.

VI. Level of Ordinary Skill in the Art

[0060] In my opinion, a person of ordinary skill in the art (“POSITA”) in the field of the ’374 patent would have had at least a bachelor’s degree in electrical engineering or related discipline, and at least two years of experience in voice or audio processing. Relevant work experience can substitute for formal education and additional education could substitute for work experience.

VII. Claim Construction

[0061] It is my understanding that in order to properly evaluate the ’374 patent, the terms of the claims must first be interpreted. It is my understanding that for the purposes of this inter partes review, the claims are

to be construed under the so-called *Phillips* standard, under which claim terms are given their ordinary and customary meaning as would have been understood by one of ordinary skill in the art in light of the specification and prosecution history, unless the inventor has set forth a special meaning for a term.

[0062] While not a departure from the customary meaning, to the extent a distinction becomes raised regarding simplex vs duplex communications, I offer the following discussion as to why a POSITA would understand the '374 patent's reference to simplex communications—particularly the claimed “simplex communication device”—to include bidirectional communication systems that limit their communications to one direction at a time.

A. “simplex communication apparatus”

[0063] Independent claim 1 recites a “simplex communication device,” which I understand dependent claims 2-8 to require as well. Dependent claims 4 and 6 each also explicitly refer to “simplex communication device.” The '374 patent does not define “simplex communication,” but it routinely describes embodiments using push-to-talk (PTT) devices a POSITA would understand uses half- or semi-duplex communications limiting bidirectional communication to one direction at a

time. '374 patent, 4:63-5:5; 8:41-64; Fig. 3. This is consistent with my experience and industry literature, such as the International Telecommunications Union (ITU) terms and definitions:

1.125 ***simplex operation***: Operating method in which transmission is made possible **alternately in each direction of a telecommunication channel**, for example, by means of manual control.²

² 1.125.1, 1.126.1 and 1.127.1 In general, duplex operation and semi-duplex operation require two frequencies in radiocommunication; simplex operation may use either one or two.

Ex. 1008 (*italics emphasis original*).

[0064] As such, consistent with the plain and ordinary meaning, the '374 patent's "simplex communication apparatus" allows communications to occur alternately in each direction of a telecommunication channel.

VIII. Claims 1-15 are Unpatentable

[0065] I have been asked to provide my opinion as to whether the Challenged Claims of the '374 patent would have been obvious in view of the prior art. The discussion below provides a detailed analysis of how the prior art references identified below teach the limitations of the Challenged Claims of the '374 patent.

[0066] As part of my analysis, I have considered the scope and content of the prior art and any differences between the alleged invention and the prior art. I describe in detail below the scope and content of the prior art, as well as any differences between the alleged invention and the prior art, on an element-by-element basis for each Challenged Claims of the '374 patent.

[0067] As described in detail below, the alleged invention of the Challenged Claims would have been obvious in view of the teachings of the identified prior art references as well as the knowledge of a POSITA.

Grounds	Claims	Basis	Prior Art
1	1-3, 5, 9, 10, 13	§ 103	IBM
2	4, 6-8, 11, 12, and 15	§ 103	IBM and Kale
3	14	§ 103	IBM, Kale, and Li

A. Ground 1: IBM Renders Claims 1-3, 5, 9, 10, and 13 Obvious

1. IBM Summary

[0068] I understand that because IBM was filed February 17, 1999, it qualifies as prior art to the '374 patent under 35 U.S.C. § 102(b). IBM discloses “systems and methods for intelligent control of microphones in speech processing.” IBM, Abstract. IBM seeks to deliver effective noise-cancellation capabilities while offering visual feedback to users about

whether the microphone system is being used effectively. IBM, [0012], [0023], [0035], [0039]; '374 patent, 24:6-18, 36:56-37:8, 37:33-47 (describing same). To this end, IBM's intelligent microphone system is "employed in speech processing applications to capture, record and preprocess audio input." IBM, [0009], [0025], Fig. 2.

[0069] IBM explains that its microphone manager 110 and associated control functions "can be designed with various frameworks depending on the target application," including "speech command and control in various environments" that commonly rely on simplex communications. IBM, [0024], [0042]. IBM also provides that its system uses speech-segment and speech-quality determinations to provide feedback to users engaging with the system. When the user activates the PTT microphone button to provide a voice command or otherwise interact with the system, IBM's system receives the speech signal to be transmitted at its microphone and transmits it to external components for recording, further processing, and providing user feedback. IBM, [0012], [0021], [0029]-[0035], [0044]-[0045], [0047], claims 7 and 17. When entering transmission mode, IBM's system also "determine[s] the boundaries of speech segments within the captured audio" to confirm the presence of speech. IBM, [0048]. Moreover, "[d]uring audio capture, the system may detect the existence of

one or more conditions that can lead to degraded acoustic quality of the captured audio.” IBM, [0047]. IBM thus determines the presence of speech in a speech signal, conducts a check on acoustic quality, and transmits the captured speech for recording, further processing, and providing notifications at the feedback output device.

[0070] IBM also describes a feedback output device (108) for providing visual notifications to a user on the processing status of microphone manager 110 for voice activity. IBM, [0025], [0036]. IBM describes varying a “lighted bar,” for example, to indicate the processing status of captured audio and whether corrections or adjustments are advisable to achieve clearer transmission of voice sounds. IBM, [0040]. In one example of varying the lighted bar, the display includes simply varying the length of brightness to turn the display on and off. IBM, [0040]. In another disclosed example, the lighted bar displays different colors on the lighted bar to notify the user of the received voice quality. IBM, [0040].

2. Claim 1

a) [1pre] A condition notification method for notifying a used condition of a simplex communication apparatus by using a light-emitting device attached to the simplex communication apparatus, the method comprising:

[0071] If the preamble is limiting, IBM discloses it by describing

advanced microphone control methods that include giving user feedback on speech-enabled devices using a “lighted bar” showing the device’s current processing status. IBM discloses intelligent microphone systems and methods “employed in speech processing applications to capture, record and preprocess audio input.” IBM, [0009], [0025], Fig. 2. IBM’s system and associated control functions “can be designed with various frameworks depending on the target application,” including “speech command and control in various environments” that use simplex communications. IBM, [0024], [0042]. IBM’s system also includes a “feedback output device” to provide visual notifications to the user on the processing status of received voice activity. IBM, [0025], [0036]. For example, IBM describes varying a “lighted bar” display to indicate the system’s processing status and whether corrections or adjustments are advisable to achieve clearer transmission of voice sounds. IBM, [0040].

b) [1a] a communication-mode determination step of determining whether a communication mode of the simplex communication apparatus is a transmission mode or a standby mode;

[0072] IBM discloses [1a] because it describes sleep-wake functionality triggered by button press of a Push to Talk (PTT) microphone. IBM explains that “the microphone manager 110 can be configured to have one or more operating modes depending on the target application,” including

environments where devices switch between sleep and wake modes. IBM, [0024], [0042]. For example, IBM describes speech-enabled applications that wake from sleep to receive and respond to vocal commands. IBM, [0024], [0042]. In particular, IBM discloses its system operates in a “detection mode”—i.e., standby mode—where the microphones deactivate and transmission of audio data for recording, speech processing, and providing feedback ceases until speech is detected. IBM, [0043]-[0044].

[0073] While '374 patent does not use the phrase “communication-mode determination step” outside of the claims, the specification explains that a PTT switch is “the communication-mode switching unit[] for switching the communication mode between the standby mode and the transmission mode.” '374 patent, 4:37-5:5, 6:6-23, 39:24-33, Fig. 1 (104), Fig. 4 (207). Likewise, IBM employs a familiar PTT switch to change communications modes. IBM, [0031]. And upon recognizing the communication mode should change based on a “speech activity event,” e.g., a PTT button press, the microphones reactivate to capture audio data. IBM, [0012], [0021], [0029]-[0035], [0044]-[0045], [0047], claims 7 and 17. Meanwhile, the transmission of audio data—i.e., transmission mode—also resumes for recording, further processing, and feedback to the user upon awakening from detection mode. IBM, [0047]-[0048], Fig. 3.

c) [1b] a sound pick-up state determination step of determining a sound pick-up state of a sound carried by a speech signal to be transmitted if the communication mode is the transmission mode;

[0074] IBM renders obvious[1b] because it determines the status of speech activity received at its microphone(s) based on both speech-segment and speech-quality determinations when transmitting feedback to the user.

Element [1f] further specifies this step “is determined based on a determination result of the speech segment determination step and an evaluation result of the speech-quality evaluation step.” ’374 patent, 43:5-8.

Likewise, the ’374 patent specification describes the “sound pick-up state determination” as the combined outcomes of speech-segment and speech-quality determinations for speech picked up by the microphone system.

The ’374 patent explains:

[F]or sound pick-up state evaluation, the sound pick-up state evaluation unit 1003 determines whether a signal input from the sound pick-up unit 1001 [i.e., a microphone] is a speech segment. Moreover, for sound pick-up state evaluation, the sound pick-up state evaluation unit 1003 evaluates the speech quality of a speech segment to be transmitted from the transmitter unit 1002. **The results of speech-segment determination and speech-quality evaluation are output to the light-emission control unit 1005**

’374 patent, 19:29-35.

[0075] In other words, the “sound pick-up state” comprises element [1d]’s speech-quality determination and element [1e.i]’s speech segment determination. Identifying speech segments and determining their associated speech quality, however, was well-known long before 2012, as was visually relaying feedback to users. *See* Section V. IBM disclosed in January 2007 that its system identifies a speaking event for processing when a user activates the PTT microphone button to provide a voice command or otherwise interacts with the system. IBM, [0012], [0021], [0029]-[0035], [0044]-[0045], [0047], claims 7 and 17. In turn, IBM’s system “will determine the boundaries of speech segments within the captured audio,” confirming the presence of speech. IBM, [0048]. IBM also explains, “[d]uring audio capture, the system may detect the existence of one or more conditions that can lead to degraded acoustic quality of the captured audio.” IBM, [0047]. IBM thus confirms the presence of speech in a speech signal, conducts a check on its acoustic quality, and transmits the captured speech for recording, further processing, and eventual use in notifications provided at the feedback output device.

d) [1c] a control step of controlling the light-emitting device so that the light-emitting device is turned off, turned on or repeatedly turned on and off based on determination results of the communication-mode determination step and the sound pick-up state determination step;

[0076] IBM discloses, or at least renders obvious, element [1c]

because one of ordinary skill would understand IBM's feedback generation system 114 turns "off," "on," or repeatedly "on and off" when it controls the lighted bar according to step [1b]. IBM, [0012], [0023], [0027], [0036], [0039], [0040], and [0060]. IBM operates with the lighted bar is turned "off" while awaiting speech activity in "detection mode" according to the communication-mode determination step. *Id.*, [0043]. But while receiving speech, IBM provides feedback according to the sound pick-up state determination. IBM, [0040], [0060]. IBM controls the lighted bar feedback in transmission mode according to speech quality and segment determinations:

[T]he system can provide feedback to a user to indicate SNR (signal to noise ratio[]), recognition confidence measures, sentence-level confidence measures, and partial playback of the recognized text as text on the screen or synthesized voice. Further, the feedback generation system (114) may track the input volume which a user speaks into a microphone and provide some form of feedback that informs the user whether the user's speech volume is too loud or too low (which can

reduce decoding accuracy) or at an optimal level for decoding. By way of specific example, the system can display a colored bar that shows the voice volume, green/ yellow/red to show normal signal, borderline, too loud, etc., and where the system can provide an indication to the user that the system is listening and capturing the speech input when a length of brightness of a lighted bar quickly tracks the user's voice activity.

IBM, [0040].

[0077] In my opinion, one of ordinary skill would appreciate IBM's feedback generation system discloses element [1c] when it varies colors and brightness to reflect sound pick-up state determinations. Indeed, IBM's lighted bar provides feedback for each of the '374 patent's described LED states corresponding to sound pick-up state determinations:

	COMMUNICATION MODE	SOUND PICK-UP STATE DETERMINATION		LED STATE
		SPEECH SEGMENT DETERMINATION	SPEECH QUALITY DETERMINATION	
No1	STANDBY MODE IBM: Detection Mode	NO DETERMINATION	NO DETERMINATION	TURN OFF
No2	TRANSMISSION MODE	NON-SPEECH SEGMENT	NO DETERMINATION	TURN ON RED
No3		SPEECH SEGMENT	BAD SPEECH QUALITY	TURN ON YELLOW
No4			GOOD SPEECH QUALITY	TURN ON AND OFF GREEN

'374 Patent, Fig. 13 (Annotated with IBM's Lighted Bar LED State)

[0078] IBM explains its feedback generation system turns on" a **red**

light when the user's voice volume and/or surrounding SNR levels have rendered the system unable to recognize speech. IBM, [0006], [0040]-[0041]. Similarly, IBM turns a **yellow** light on or "repeatedly on and off" when borderline readings indicate poor speech quality as the system tracks voice activity. Id. Finally, IBM turns a **green** light "**repeatedly on and off**" when identifying good quality speech while "quickly track[ing] the user's voice activity" by brightness level. Id.

[0079] The '374 patent also makes clear the claimed LED on-off sequences were never considered inventive. '374 patent, 39:50 ("The light-emission control unit controls the light-emitting device (LED) in the following manner in the embodiments, however, **the present invention is not limited thereto.**") (emphasis added). The '374 patent instead explains that other alternatives for conveying the same information work just as well. Id., 39:61-40:2; *see also* Exs. 1009-1013. Accordingly, a POSITA would have understood IBM discloses, or at least renders obvious, found element [1c] obvious over based on IBM's feedback generation system and use of a versatile lighted bar, especially given the ubiquity of using on-off sequences to provide user feedback.

e) [1d] a speech-quality evaluation step of evaluating speech quality of the speech signal to be transmitted,

[0080] IBM discloses [1d] because, during audio capture, IBM

seeks to evaluate one or more conditions leading to degraded acoustic quality of the captured audio. IBM, [0047]. As discussed with respect to [1c], IBM controls feedback displayed on the lighted bar reflecting a speech quality determination (discussed here) and speech segment determination (discussed below with respect to [1e.i]) as part of determining the sound pick-up state (element [1b]). IBM explains that its acoustic quality control methods “may be implemented using any one or more of known speech/audio processing techniques that are suitable for microphone/speech processing applications.” IBM, [0060]. In one embodiment, IBM describes this step performed by an “acoustic quality control system [that] detects conditions regarding capturing of audio input to the microphone system.” IBM, [0012]-[0013], [0023].

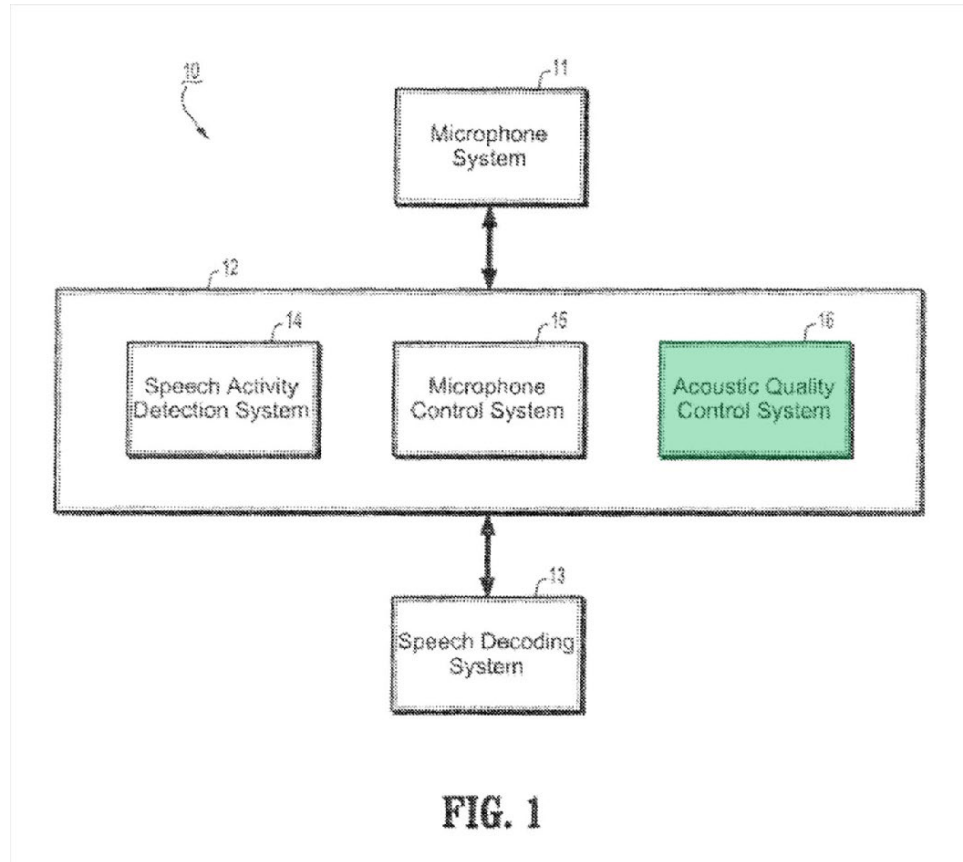


Fig. 1 (Annotated)

[0081] In other embodiments, IBM describes “[t]hese acoustic quality control functions [being] supported by system modules such as the environment interpretation system (113) and feedback generation system (114).” IBM, [0036]. “In particular, the environment interpretation system (113) can implement methods for processing sensor data and environmental static noise to detect conditions that could possibly have an adverse affect [sic] on the decoding accuracy of speech data” IBM, [0037]. This allows “the environment interpretation system (113) [to] determine or otherwise detect the static environmental noise in an audio signal received

via a user microphone” IBM, [0037]. Thus, IBM captures and monitors audio data to track noise levels affecting speech quality by comparing it to the captured speech. IBM, [0038], [0052].

[0082] IBM also explicitly identifies other means for making speech quality assessments on received speech. For example, IBM identifies and provides feedback regarding audio quality based on volume input level. IBM, [0040]. In particular, IBM provides “some form of feedback that informs the user whether the user's speech volume is too loud or too low (which can reduce decoding accuracy) or at an optimal level for decoding.” IBM, [0040]. Furthermore, the environmental interpretation system (113) can also process sensor data to “detect [the] location and distance between a microphone and a speaker's mouth,” allowing the system to determine the microphone’s positioning has caused degradation of the audio input. IBM, [0040]. IBM’s system then provides feedback indicating the poor quality to “allow the user to make manual adjustments.” IBM, [0023], [0040].

[0083] IBM also seeks to automatically improve audio quality before providing feedback. IBM, [0036]; ’374 patent, 40:3-14 (same). IBM’s system evaluates captured audio input for “acoustic characteristics that adversely affect a decoding accuracy” of the speech signal to be transmitted. IBM, [0026]. As explained for [1c], IBM’s feedback generation system

(114) implements “methods for processing and rendering feedback data output from the environment interpretation system (113), which is presented to a user via the feedback output device (108).” IBM, [0039], [0012] (“In one embodiment, the acoustic quality control system provides feedback to a user regarding a detected condition having an adverse affect [sic] on an acoustic quality of audio input.”), [0023], [0036].

f) [1e] wherein, in the sound pick-up state determination step, the sound pick-up state of the sound carried by the speech signal to be transmitted is determined if the communication mode is the transmission mode based on the speech quality of the speech signal evaluated by the speech-quality evaluation step; and

[0084] IBM discloses [1e] because it determines the speech quality of captured speech when, e.g., a user presses the microphone button to bring the system out of detection mode and into transmission mode. As discussed for [1a], “microphone manager 110 can be configured to have one or more operating modes depending on the target application.” IBM, [0042]. This includes embodiments where microphone manager 110 waits in “detection mode” for the onset of speech at a PTT microphone, deactivating the microphone and ceasing transmission of audio data for recording, speech processing, and feedback while waiting. IBM, [0043]-[0044]. But once the microphone manager 110 detects speech (e.g., by microphone button press),

it reactivates the microphones to capture the corresponding audio data and resumes transmitting upon waking from detection mode. IBM, [0047]-[0048], Fig. 3. And while the microphone remains activated, and audio input to the microphone is captured and recorded, microphone manager 110 evaluates the audio input for acoustic quality. IBM, [0012], [0039]-[0040], [0045], claim 2.

(1) [1e.i] a speech-segment determination step of determining whether or not the speech signal to be transmitted is a speech segment if the communication mode is the transmission mode,

[0085] IBM discloses [1e.i] because it discloses a variety of methods for distinguishing between speech and nonspeech segments as part of the feedback presented to the user. IBM discloses, for example, that microphone control system (15) operates with speech activity detection system (14) to identify segments of captured audio having relevant speech data. IBM, [0021]. In particular, “the speech activity detection system (14) implements functions to detect one or more types of speech activity events during audio capturing, which serve as indicators of speech input to the microphone system (11).” IBM, [0021]. Speech activity events include, e.g., user interactions with microphone switch buttons or “speech intent events” indicating a user's intention to speak. IBM, [0021]. Speech intent events can also be based on biometric feature data, behavioral characteristics, and/or

other types of events. IBM, [0021].

[0086] As IBM explains, detecting speech activity events provides the system with information for determining the timing of speech from a user, e.g., speech onset, speech offset, ongoing speech, etc. IBM, [0022]. The microphone control system (15) uses these speech activity events to determine, among other things, boundaries of speech segments within captured audio for selective sending to the speech decoding system (13). IBM, [0022]. In other embodiments, IBM uses the microphone interaction control system (112) to identify boundaries of speech segments within captured audio based on the timing of speech. IBM, [0021], [0027], [0028] (“The microphone interaction control system (112) uses timing of speech activity events to . . . determine the boundaries of speech segments within captured audio.”)

[0087] In still other embodiments, IBM describes using control methods based on multimodal input from sensor devices used to control the audio capture and determine speech segments within it. IBM, [0026], [0048]. For example, user behavior interpretation system (115) and audio data processing system (116) also implement methods supporting speech activity detection. IBM, [0029]. User behavior interpretation system (115) collects and processes multimodal input from various sensors (e.g., cameras,

biometric sensors, tactile sensors, etc.) as a user interacts with the system to determine the timing of speech activity (speech onset/offset, etc.). IBM, [0029]. User behavior interpretation system (115), on the other hand, processes image data of the speaker to extract facial features (e.g., lips, mouth, pose) or other body data to detect “speech onset to activate capturing and recording of audio input based on visual speech clues.” IBM, [0034]. Furthermore, audio data processing system (116) employs known audio-based techniques for processing the audio input to detect onset of speech. IBM, [0029]. For example, IBM explains that audio data processing system (116) identifies the onset of speech when the speech energy of the audio input received crosses a predetermined threshold. IBM, [0029]. Finally, IBM explains the microphone interaction control system (112) can also detect speech onset by considering the timing of speech in conjunction with the above-discussed information collected by other system components. IBM, [0029]-[0031].

g) [1f] wherein, in the sound pick-up state determination step, the sound pick-up state of the sound carried by the speech signal to be transmitted, if the communication mode is the transmission mode, is determined based on a determination result of the speech-segment determination step and an evaluation result of the speech-quality evaluation step.

[0088] IBM discloses, or at least renders obvious, element [1f]

obvious because its system determines the status of speech activity received at its microphone(s) based on both the presence of speech and assessed quality. Moreover, in considering IBM, one of ordinary skill would have appreciated how to apply the speech processing techniques already known to the industry in 2012, namely identifying and evaluating speech.

[0089] More than five years before the '374 patent's earliest priority date, IBM disclosed a system using known speech-segment and speech-quality techniques to provide feedback to users on captured audio and encourage proper use of its microphone. As explained, IBM's system determines both "the boundaries of speech segments within the captured audio" and "detect[s] the existence of one or more conditions that can lead to degraded acoustic quality" while in transmission mode. IBM, [0047]. The system then generates visual feedback for the user by controlling a "lighted bar" display to indicate different outcomes of the speech and quality determinations. IBM, [0023], [0036], [0040]. For example, IBM explains that its system provides feedback regarding user's voice volume (quality determination), SNR (quality determination), word and sentence confidence measures (segment determination), partial playback (segment determination), etc. IBM, [0023], [0036], [0040]. IBM also describes "quickly track[ing] the user's voice activity" by brightness level (segment

determination). IBM, [0006], [0040]-[0041]. Consistent with its shared aim with the '374 patent, IBM explains that this feedback “allow[s] the user to make manual adjustments or take appropriate measures to ensure the acoustic quality of the captured audio,” such as adjusting the microphone, quieting the room, speaking more attentively towards the microphone, or just speaking louder. IBM, [0023], [0036], [0040]; see also '347 patent, 1:29-42.

h) Claim 2: The condition notification method according to claim 1, wherein, in the control step, the light-emitting device is turned on if the communication mode is the transmission mode, whereas the light-emitting device is repeatedly turned on and off based on the sound pick-up state determined in the sound pick-up state determination step if the communication mode is the transmission mode.

[0090] IBM also renders dependent claim 2 obvious because it discloses a lighted bar with several ways of conveying sound pick-up state information on captured speech and assessed audio quality. IBM's lighted bar provides feedback in multiple ways reflecting the system's processing condition via a lighted display while receiving and processing voice commands. IBM, [0040]. IBM discloses a feedback generation system that turns on a red, yellow, or green light in transmission mode, and repeatedly turns on and off when the system can identify and track speech activity:

	COMMUNICATION MODE	SOUND PICK-UP STATE DETERMINATION		LED STATE
		SPEECH SEGMENT DETERMINATION	SPEECH QUALITY DETERMINATION	
No1	STANDBY MODE IBM: Detection Mode	NO DETERMINATION	NO DETERMINATION	TURN OFF
No2	TRANSMISSION MODE	NON-SPEECH SEGMENT	NO DETERMINATION	TURN ON RED
No3		SPEECH SEGMENT	BAD SPEECH QUALITY	TURN ON YELLOW
No4			GOOD SPEECH QUALITY	TURN ON AND OFF GREEN

'374 Patent, Fig. 13 (Annotated with IBM Lighted Bar LED State)

[0091] As discussed for element 1[c], IBM turns on a red light when the system cannot recognize speech due to, e.g., the user's voice volume or surrounding SNR levels being too high. IBM, [0006], [0040]-[0041]. IBM also turns on a yellow light to indicate borderline quality readings for captured speech. Id. And when receiving good quality speech, IBM turns its lighted bar green and "repeatedly on and off" while "quickly track[ing] the user's voice activity" by brightness level. Id.

[0092] Moreover, one of skill would recognize the conventionality of providing feedback for microphone systems using a simple "on" and "off" display sequence. *See Technology Background*; Exs. 1009-1013. As the '374 patent explains, the claimed on-off sequences for a given LED state are not a

point of novelty. '374 patent, 39:50 (explaining that the “the present invention is not limited thereto.”).

3. Claim 3: The condition notification method according to claim 1, wherein, in the speech-quality evaluation step, the speech quality of the speech signal to be transmitted is evaluated based on a volume level of a sound to be picked up.

[0093] IBM discloses claim 3 because it describes monitoring speech volume and providing feedback to the speaker about whether they are speaking too loud or too low. IBM explains that its acoustic quality control methods “may be implemented using any one or more of known speech/audio processing techniques that are suitable for microphone/speech processing applications.” IBM, [0047], [0060]. Moreover, disclosed embodiments evaluate captured audio input for “acoustic characteristics that adversely affect a decoding accuracy” of the speech signal to be transmitted. IBM, [0026]. As explained for [1c], IBM also describes several different types of feedback available from the feedback generation system. [0040], [0060]. For example, IBM describes providing feedback on speech quality evaluated based on the volume level of the speaker:

Further, the feedback generation system (114) may track the input volume which a user speaks into a microphone and provide some form of feedback that informs the user whether the user's speech volume is too loud or too low (which can reduce decoding

accuracy) or at an optimal level for decoding. By way of specific example, the system can display a colored bar that shows the voice volume, green/ yellow/red to show normal signal, borderline, too loud, etc., and where the system can provide an indication to the user that the system is listening and capturing the speech input when a length of brightness of a lighted bar quickly tracks the user's voice activity.

IBM, [0040].

4. Claim 5

a) [5a] The condition notification method according to claim 1, wherein in the speech-segment determination step it is determined whether or not the speech signal to be transmitted is a speech segment if the communication mode is the transmission mode, wherein,

[0094] IBM also discloses [5a], as I discussed with respect to [1e.i] as it relates to “a speech-segment determination step of determining whether or not the speech signal to be transmitted is a speech segment if the communication mode is the transmission mode.”). As discussed, “microphone manager 110 can be configured to have one or more operating modes depending on the target application.” IBM, [0042]. This includes applications where microphone manager 110 waits in “detection mode” for speech activity and switches to transmission mode at the onset of speech (e.g., based the user pressing the talk button a PTT microphone). IBM,

[0043]-[0048], Fig. 3. While the microphone remains activated, microphone manager 110 captures, records (e.g., in audio recording/storage device 107), and evaluates the audio input for speech segments as part of providing feedback to the user. IBM, [0021]-[0022], [0026]-[0031], [0034], [0042]-[0048].

b) [5b] in the speech-quality evaluation step, the speech quality of the speech signal to be transmitted is evaluated while the speech signal to be transmitted is being determined as the speech segment.

[0095] IBM also discloses [5b] because it discloses contemporaneous speech segment and quality determinations. Upon exiting “detection mode,” and entering transmission mode, microphone interaction control system (112) resumes identifying speech segments to selectively feed audio data (i.e., isolated speech) to automatic speech recognition (ASR) system (120). IBM, [0030]. Meanwhile, IBM’s acoustic quality controls systems (e.g., environment interpretation system 113) “function[] to detect certain conditions that could potentially reduce the decoding accuracy (by the ASR system (120)) of speech input.” IBM, [0036]. In other words, one of skill would understand IBM’s microphone manager 110 assesses speech quality conditions as it prescreens speech segments for ASR 120. IBM, [0023], [0037].

5. Independent Claim 9

a) [9pre] A communication apparatus comprising:

[0096] If the preamble is limiting, IBM discloses [9pre] for the reasons I discussed with respect to [1pre]. IBM discloses an intelligent microphone control system “employed in speech processing applications to capture, record and preprocess audio input in a way that optimizes decoding accuracy of spoken input.” IBM, [0008]. IBM also explains that its microphone manager 110 “can be designed with various frameworks depending on the target application,” including “speech command and control in various environments.” IBM, [0024], [0042], Figs. 1-2.

a) [9a] a first pick-up unit configured to pick up a voice sound;

[0097] IBM discloses [9a] because it uses a microphone to receive speech. IBM discloses systems and methods for the “intelligent control of microphones.” IBM, Title; *see also* ’374 patent, 4:7 (“The sound pick-up unit 101 is a microphone”). IBM describes a microphone manager system 110 to receive audio captured by microphone system (101) (a first pick-up unit) for recording, speech processing, and providing feedback to the user about the current speech capture. IBM, [0043]- [0048], Figs. 1-2.

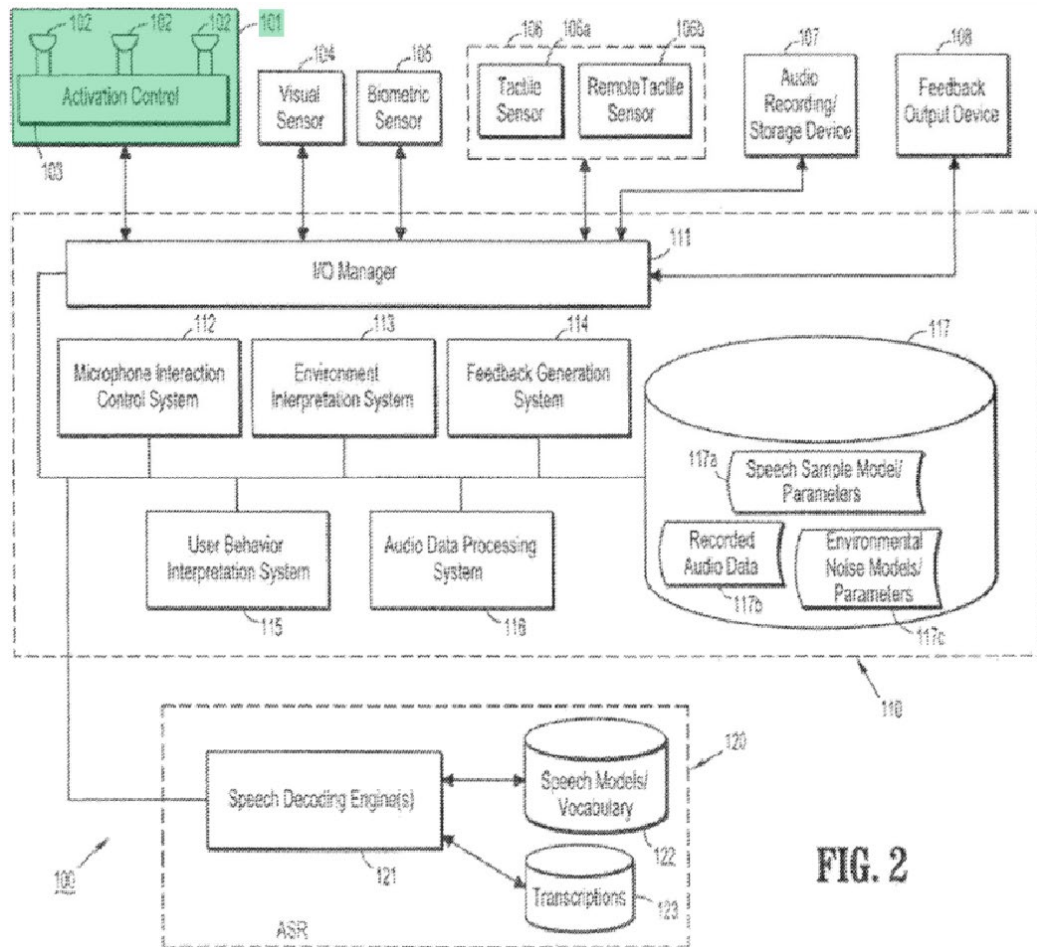


FIG. 2

IBM Figure 2 (Microphone System 101)

b) [9b] a transmitter unit configured to transmit the voice sound picked up by the first pick-up unit to outside as a first speech signal;

[0098] IBM discloses [9b] because it discloses an I/O manager for communicating with internal and external components associated with microphone manager 110's operation. For example, IBM's "I/O manager module (111) performs various functions such as managing the I/O data that is input and output to the various components (101, 104, 105, 106, 107, 108) and otherwise providing an API for communication of messages and data

between the microphone manager 110 the external systems/components.”

IBM, [0025], [0027]. One of skill would appreciate IBM’s I/O manager includes a transmitter for transmitting data among internal and external system components. As I discussed regarding [1c] and [1e.i], IBM discloses I/O manager transmitting captured audio received by its microphone system 101 as a signal to, e.g., internal and external components to microphone manager 110, such as ASR 120, Audio Recording/Storage Device 107, and Feedback Output Device 108.

c) [9c] a communication-mode switching unit configured to switch a communication mode between a standby mode in which the transmitter unit does not transmit the speech signal and a transmission mode in which the transmitter unit transmits the speech signal;

[0099] For the reasons that I also discuss for [1a], IBM discloses [9c] because it describes using well-known PTT switches as one means to switch between communications modes. IBM, [0031]. This is also what the ’374 patent describes. ’374 patent, 4:37-5:5, 6:6-23, 34:14-27, 39:24-33, Fig. 1 (104), Fig. 4 (207). For example, IBM provides embodiments operating in speech command and control environments, which often switch between wake and sleep states to receive and respond to user commands. IBM, [0042]. IBM’s system, for example, deactivates microphones and ceases transmission of audio data for recording, speech processing, and

feedback while in “detection mode”—i.e., standby mode. IBM, [0043]-[0044].

[0100] IBM’s system remains in detection mode until it identifies speech activity, switching to transmission mode with the onset of speech. *Id.*, [0043]-[0044]. IBM describes several ways to identify a user’s intention to speak, including traditional PTT microphones or sensed conditions. *Id.*, [0043]-[0046], [0020] (“[S]peech activity events include user interactions with microphone switch buttons, speech intent events that indicate a user’s intention to speak into the microphone (as determined based on processing of biometric feature data of the user and/or from data acquired from tracking behavioral characteristics of the user, etc.)” For target applications not reliant on PTT for speech detection, “the microphone is maintained ‘deactivated’ by the microphone manager (110) in the sense that the microphone is ‘on’ and the system is listening for speech input, but that the audio input is not being recorded or otherwise applied to the speech decoding system until detection of a speech activity event (e.g., speech onset) (step 201).”). *Id.*

[0101] But upon recognizing a “speech activity event,” e.g., a PTT button press, the microphones activate to capture audio data. IBM, [0012], [0021], [0029]-[0035], [0044]-[0045], [0047], claims 7 and 17. Meanwhile,

the transmission of audio data—i.e., transmission mode—also resumes for recording, further processing, and feedback to the user upon awakening from detection mode. IBM, [0047]-[0048], Fig. 3.

d) [9d] a sound pick-up state determination unit configured to determine a pick-up state of the voice sound picked up by the first pick-up unit;

[0102] For the reasons I further discuss for [1b], IBM renders [9d] obvious because it describes compiling the results of speech-segment and speech-quality determinations as part of providing noise-cancellation on received speech and providing the user feedback on a lighted bar.

e) [9e] a light emission device configured to emit light;

[0103] IBM discloses [9e] because it discloses a feedback output device, including “lighted bar” for displaying feedback to users. IBM, [0040]; *see also* ’374 patent, 26:2 (describing the light-emitting device as “for visual notification”), 26: 51-52, 33:17-19.

[0104] The ’374 explains that “light-emitting device 1210 [is] for visual notification,” including to “notify a user of speech quality” and “on correct holding of the main microphone.” ’374 patent, 26:2, 26: 51-52, 33:17-19. In the same way, IBM discloses a “lighted bar” for providing feedback regarding speech quality and the correct holding of the microphone. IBM, [0012] (“[T]he acoustic quality control system provides

feedback to a user regarding a detected condition having an adverse affect [sic] on an acoustic quality of audio input.”), [0023] (explaining feedback “allow[s] the user to make manual adjustments or take appropriate measures to ensure the acoustic quality of the captured audio.”).

[0105] In one embodiment, IBM explains that its system provides feedback regarding the user’s voice volume and detected signal-to-noise ratio (SNR). IBM, [0040]. IBM explains, for example, that “the system can provide an indication to the user that the system is listening and capturing the speech input when a length of brightness of a lighted bar quickly tracks the user's voice activity.” IBM, [0040]. Moreover, IBM further discloses turning the lighted bar on to notify the user of voice volume quality using different colors, e.g., green, yellow, or red to indicate a normal, borderline, or overly loud signal. IBM, [0040].

f) [9f] a control unit configured to control the light-emitting device so that the light-emitting device is turned off, turned on or repeatedly turned on and off based on the communication mode switched by the communication-mode switching unit, and the pick-up state of the voice sound picked up by the first pick-up unit determined by the sound pick-up state determination unit; and

[0106] As I discussed with respect to [1c]’s “control step,” IBM renders [9f] obvious because one of ordinary skill would understand IBM’s feedback generation system turns its lighted bar “off,” “on,” or repeatedly

“on and off” when it controls the feedback displayed to a user regarding the sound-pick up state for captured speech determined at element [9d].

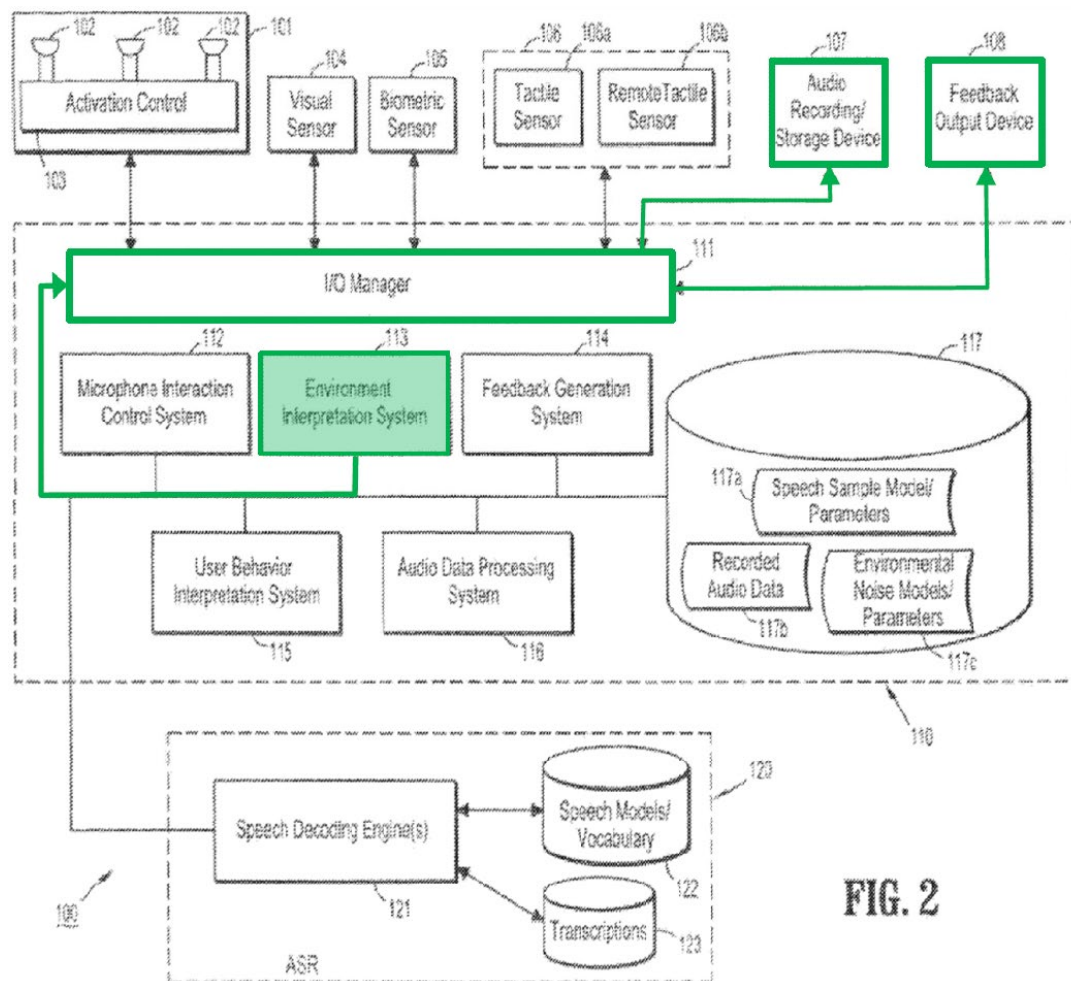
g) [9g] a speech-quality evaluation unit configured to evaluate speech quality of the first speech signal to be transmitted by the transmitter unit,

[0107] IBM also discloses [9g] because IBM detects conditions leading to degraded acoustic quality in the captured audio transmitted for feedback to the user. IBM, [0047]. IBM explains, for example, “[t]he acoustic quality control system detects conditions regarding capturing of audio input to the microphone system.” IBM, [0012]-[0013], [0023].

[0108] In other embodiments, IBM describes “[t]hese acoustic quality control functions [being] supported by system modules such as the environment interpretation system (113)” IBM, [0036]. “In particular, the environment interpretation system (113) can implement methods for processing sensor data and environmental static noise to detect conditions that could possibly have an adverse affect [sic] on the decoding accuracy of speech data” IBM, [0037].

[0109] As I explained regarding [1d], environment interpretation system (113) evaluates the audio signal for speech quality based on static noise detected in the signal, distance or orientation between the microphone and speakers mouth, and/or other data indicating an adverse effect on

acoustic quality of the audio input. IBM, [0026], [0037]. IBM's feedback generation system (114) implements "process[es] and render[s] feedback data output from the environment interpretation system (113)," which gets transferred to external components via I/O Manager (111) for presentation to a user via the feedback output device (108):



IBM, Fig. 2 (Environmental Interpretation System 113 Feedback)

IBM, Fig. 2., [0039], [0012] ("In one embodiment, the acoustic quality control system provides feedback to a user regarding a detected condition

having an adverse affect [sic] on an acoustic quality of audio input.”),
[0023], [0036].

h) [9h] wherein the sound pick-up state determination unit determines the sound pick-up state of the voice sound picked up by the first sound pick-up unit based on the speech quality of the speech signal evaluated by the speech-quality evaluation unit; and

[0110] IBM discloses [9e] because, as discussed regarding [1e]’s speech-quality evaluation step, microphone manager (110) determines the speech quality of speech captured when, e.g., a user presses the microphone button to bring the system out of detection mode and into transmission mode.

i) [9i] a speech-segment determination unit configured to determine whether or not the first speech signal to be transmitted by the transmitter unit is a speech segment,

[0111] IBM discloses [9i] because, as I discussed regarding [1e.i]’s “speech-segment determination step,” microphone manager (110) employs multiple systems to determine speech segments on captured audio according to discussed pre-processing and speech processing techniques. For example, microphone control system (15) operates in conjunction with speech activity detection system (14) to identify segments of captured audio having relevant speech data. IBM, [0021]. Moreover, microphone interaction control system

(112) uses the timing of speech to identify boundaries of speech segments, IBM, [0021], [0027], and user behavior interpretation system (115) and audio data processing system (116) further provide speech activity and segment detection support based on multimodal input from various sensors (e.g., cameras, biometric sensors, tactile sensors, etc.). IBM, [0029].

j) [9i] wherein, the sound pick-up state determination unit determines the sound pick-up state of the sound to be transmitted as the first speech signal based on a determination result at the speech-segment determination unit and an evaluation result at the speech-quality evaluation unit.

[0112] For the reasons I further detail as it relates to Claim 1's corresponding [1f], IBM renders [9i] obvious because IBM's system determines the status of speech activity received at its microphone(s) based on both the presence of speech and assessed quality.

6. Claim 10: The communication apparatus according to claim 9, wherein the light-emitting device is turned on if the communication mode is switched into the transmission mode, whereas the light-emitting device is repeatedly turned on and off based on the pick-up state determined by the sound pick-up state determination unit if the communication mode is switched into the transmission mode.

[0113] IBM also renders dependent claim 10 obvious because, as I further discuss as to corresponding dependent claim 2, IBM discloses a versatile, lighted bar for providing feedback in a number of ways to reflect

the system's processing condition through controlling light emission while processing received voice commands. IBM, [0040]. One of ordinary skill would appreciate from IBM that its feedback generation system discloses a lighted bar that is turned off, turned on, or repeatedly turned on and off according to speech segment and speech quality determinations based on its communication mode.

7. Claim 13

a) [13a] The communication apparatus according to claim 9, wherein the control unit controls the light-emitting device so that the light-emitting device is turned off if the communication mode is the standby mode,

[0114] IBM renders [13a] obvious because IBM's system does not illuminate —i.e., turns off—its lighted bar when operating in “detection mode.” *See* IBM, [0040]-[0044]. IBM discloses different operating modes “depending on the target application,” including speech command and control environments that switch between sleep and wake states to receive and respond to a user's commands. IBM, [0042]. In detection mode (standby mode), IBM's system does not provide feedback to users using the lighted bar and is “turned off.” *Id.*, [0040], [0043]-[0044]; *see also* '374 patent, 5:25-27 (“Thus, the light-emission control unit 105 turns off the light-emitting device 106 with no drive current supplied thereto.”). In particular,

IBM discloses operating in “detection mode”—standby mode—by deactivating the microphones and ceasing transmission of audio data for recording, speech processing, and feedback . IBM, [0043]-[0044]. IBM’s system remains in detection mode until detecting the onset of speech activity based on, e.g., “user activation of a microphone control button” or sensed conditions (e.g., visual or biometric sensing). *Id.*, [0043]-[0046], [0020] (“[S]peech activity events include user interactions with microphone switch buttons, speech intent events that indicate a user's intention to speak into the microphone (as determined based on processing of biometric feature data of the user and/or from data acquired from tracking behavioral characteristics of the user, etc.”).

b) [13b] the light-emitting device is turned on if the communication mode is the transmission mode and if the speech signal to be transmitted is not the speech segment, the light-emitting device is turned on if the communication mode is the transmission mode, if the speech signal to be transmitted is the speech segment, and if the speech quality of the speech signal to be transmitted is evaluated as bad, and the light-emitting device is repeatedly turned on and off if the communication mode is the transmission mode, if the speech signal to be transmitted is the speech segment, and if the speech quality of the speech signal to be transmitted is evaluated as good.

[0115] IBM renders [13b] obvious because it discloses varying the brightness of its lighted bar to track voice activity and indicating voice

quality using different colors, as further detailed in discussing [1c]’s “control step of controlling the light-emitting device so that the light-emitting device is turned off, turned on or repeatedly turned on and off based on determination results of the communication-mode determination step and the sound pick-up state determination step.”

B. Ground 2: IBM and Kale Render Obvious Claims 4, 6-8, 11, 12, and 15

1. Kale Summary

[0116] I understand that because Kale was filed July 28, 2010, and published February 2, 2012, it qualifies as prior art to the ’374 patent under at least §102(a). Kale discloses a “noise cancellation device” and method. The device includes inter alia two microphones, a direction of arrival (DOA) module, a beamformer, and an Inverse Fast Fourier Transform (IFFT) module. Kale, [0028], Fig. 8. Kale’s microphones detect “two [sound] sources,” (e.g., “speech” and “noise”) and “generate respective raw audio data 835, 845.” Kale, [0028] [0043], FIGS. 8-9.

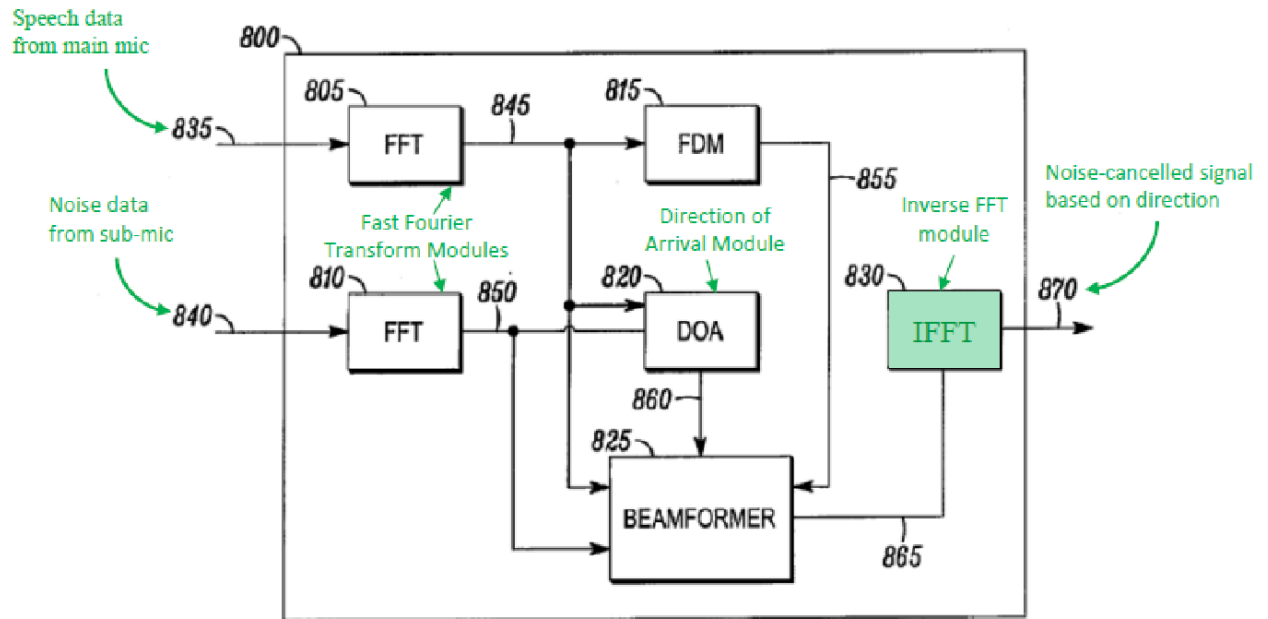


FIG. 8

Kale, Fig. 8 (Annotated)

[0117] As reflected in Figure 8, the FFT modules “determine which of the raw audio data signals comprises the speech signal” using “factors ... such as common formants, formants with known patterns, etc.” and output a frequency-domain “speech signal” 845, 850 to the DOA module and beamformer. Kale, [0003], [0043]-[0044], FIG. 8. The DOA module “determin[es] the direction from which the speech data arrived” and outputs a “DOA data signal” 860 containing the direction information to the beamformer. Kale, [0028], [0031], [0046], FIGS. 8-9. Finally, the beamformer uses the inputted speech signal and direction information to “modify [the] audio signal to isolate a speech therein” or “enhance a speech of an audio signal” by “eliminating” the noise contained in the received

audio signal” before applying a an inverse FFT to return to the time domain. Kale, [0040], [0047]-[0048], FIGS. 8-9.

2. Reasons to Combine IBM and Kale (Ground 2)

[0118] In my opinion, IBM and Kale are analogous art to each other and the '374 patent because they are in the same field of endeavor: audio signal processing. Kale, Abstract; IBM, [0002]-[0008], [0025]; '374 patent, 1:17-2:35, 38: 3-13. IBM and Kale also share a common goal with the '374 patent: improving noise reduction methods to enhance voice sound quality. Kale, Abstract; IBM, [0012], [0023], [0035], [0039]; '374 patent, 24:6-18, 36:56-37:8, 37:33-47. Kale's disclosure merely provides additional techniques for doing so based on microphone direction.

[0119] Kale also makes clear its noise cancellation device may serve as a component to a larger system, explaining the device “may be incorporated with any electronic device that includes an audio receiving device such as a microphone.” Kale, [0028]. In my opinion, one of ordinary skill would have had a reasonable expectation of success in configuring IBM's intelligent microphone system with Kale's voice direction detector and associated noise-processing techniques because—as Kale suggests—it would have merely required routine skill in the art to do so. For example, a POSITA would have understood Kale's noise-suppression techniques based

on voice-direction would precede IBM's existing noise reduction techniques to further enhance voice sound quality.

[0120] Given that IBM and Kale disclose similar systems that operate in a similar manner and are directed to similar problems, it is my opinion a POSITA would have been motivated to consider their combined teachings. But for specific claim elements, I may provide additional rationales for the combination.

3. Claims 4 and 11

a) Claim 4

(1) [4a] The condition notification method according to claim 1 further comprising a noise cancellation step of performing noise cancellation to a first speech signal input from a first sound pick-up unit by using a second speech signal input from a second sound pick-up unit, the first and second sound pick-up units being provided for the simplex communication apparatus,

[0121] In my opinion, IBM renders 4[a] obvious, taken alone or in combination with Kale, because both seek to improve acoustic quality by cancelling noise from a speech signal received at one microphone based on noise detected by another microphone. In fact, the '374 patent's "BACKGROUND OF THE INVENTION" section recognizes the claimed dual microphone setup was already well-known at the time of filing. '374

patent, 1:29-32.

[0122] Regardless, IBM discloses a microphone system with “a microphone that is dedicated to capturing environmental noise and another microphone for capturing speech input.” IBM, [0037], [0053]. IBM’s environmental interpretation system (113) “implement[s] known techniques for determining whether noise captured by the microphone is speech or static environmental noise.” IBM, [0053]. For example, IBM explains:

[E]nvironment interpretation system (113) can determine or otherwise detect the static environmental noise in an audio signal received via a user microphone or a microphone specifically implemented as a detector of the non-speech environment and generate and update models/parameters of the environmental noise (117c), which are stored in local memory (117).

IBM, [0037], [0038], [0043].

[0123] IBM further describes making “adjustments to microphone and/or decoding parameters” to reduce environmental noise levels based on the audio signals received. IBM, [0037]. For example, IBM describes “transforming a speaker's voice by adapting microphone volume sensitivity depending on environmental input received from the detector of non-speech environment.” IBM, [0037]. Based on these disclosures, in my opinion, one of ordinary skill would understand IBM’s adapting microphone sensitivity discloses [4a] because it performs noise cancellation by manipulating a main

signal (from the voice microphone) based on another signal (from the noise microphone). Said another way, IBM effects noise cancellation to the speech signal received from a microphone dedicated for speech based on the noise observed from the noise signal received from the microphone dedicated to noise.

[0124] And to the extent it becomes argued adjusting microphone sensitivity does not render [4a] obvious, Kale further discloses a “noise cancellation device” that performs known techniques for noise cancellation using multiple microphones that also render [4a] obvious. Kale, [0003], [0028], FIG. 8. Kale’s noise cancellation device provides “two microphones” for detecting “two [sound] sources,” e.g., “speech” and “noise.” Kale, [0028] [0043], FIGS. 8-9. Using the speech and noise sources, Kale describes “modify(ing) an audio signal to isolate a speech therein” by “eliminating” noise contained in the received audio signal. Kale, [0003], [0034], Abstract. In other words, Kale eliminates noise in the dedicated speech signal by removing parts of its audio profiles common to the noise detected by the noise microphone. Kale, Abstract, [0003], [0034], [0040], [0047]-[0048], FIGS. 8 and 9.

(2) [4b] wherein, speech quality of a speech signal after the noise cancellation is evaluated in the speech-quality evaluation step.

[0125] IBM discloses [4b] because, like the '374 patent, it describes embodiments that seek to improve audio quality automatically before evaluating the captured audio. IBM, [0010], [0012], [0018], [0023], [0027], [0036], [0039], and [0040]; '374 patent, 40:3-14. As an initial matter, one of ordinary skill would readily appreciate that a control system would, in the ordinary course of events, evaluate quality before implementing controls that adjust for better quality. But as IBM itself explains, "[b]y automatically adjusting the microphone parameters, the system may be better adapted to the speaker under changing noise levels to thereby enhance the quality of the captured speech data." IBM, [0054]. In other words, IBM seeks to enhance the quality of its received audio before evaluating and providing feedback to the user on the audio's quality. In my opinion, one of ordinary skill would understand from IBM's disclosure that it evaluates captured audio for quality after applying noise-cancellation to improve its quality.

b) Claim 11

(1) [11a] The communication apparatus according to claim 9 further comprising a second sound pick-up unit configured to pick up a voice sound;

[0126] IBM discloses [11a] because it discloses using more than

one microphone to pick up speech. IBM's microphone system 101 includes "a plurality of different microphones (102) that are implemented for different tasks." IBM, [0038]. "Moreover, one or more microphones (102) can serve to capture speech input from speakers, and one or more of the microphones (102) may [be] used to monitor/capture environmental audio data, which is used to track changes in environmental noise levels." IBM, [0038], [0046]. In particular, "the microphone system (101) in FIG. 2, for example, can include a microphone that is dedicated to . . . capturing speech input." IBM, [0053], Fig. 2.

(2) [11b] a noise cancellation unit configured to perform noise cancellation to the first speech signal input from the first sound pick-up unit by using a second speech signal input from the second sound pick-up unit,

[0127] In my opinion, IBM, taken alone or in combination with Kale, renders [11b] obvious for the reasons explained regarding [4a]'s: "performing noise cancellation to a first speech signal input from a first sound pick-up unit by using a second speech signal input from a second sound pick-up unit, the first and second sound pick-up units being provided for the simplex communication apparatus."

(3)[11c] wherein the transmitter unit transmits a speech signal after the noise cancellation by the noise cancellation unit and the speech-quality evaluation unit evaluates speech quality of the speech signal after the noise cancellation.

[0128] IBM discloses [11c] through at least its embodiments seeking to improve audio quality automatically prior to evaluating captured audio. IBM, [0012], [0023], [0027], [0036], [0039], and [0040]. “By automatically adjusting the microphone parameters, the system may be better adapted to the speaker under changing noise levels to thereby enhance the quality of the captured speech data.” IBM, [0054]; *see also* my discussion as to element [4a]; Kale, [0028] [0043], FIGS. 8-9. After noise cancellation, “feedback generation system (114) may implement methods to provide feedback to a user about different parameters/conditions affecting the quality of captured audio due to inefficient microphone use or sub-optimal decoding parameters, etc., *which could not be fixed automatically.*” IBM, [0039].

4. Claim 7

(1)[7a] The condition notification method according to claim 4, wherein, in the control step, the light-emitting device is turned off if the communication mode is the standby mode,

[0129] IBM renders [7a] obvious because IBM’s system does not illuminate its lighted bar when operating in “detection mode.” IBM discloses

different operating modes “depending on the target application,” including speech command and control environments that switch between sleep and wake states to receive and respond to a user’s commands. IBM, [0042]. In detection mode, IBM’s system does not provide feedback to users using the lighted bar and is “turned off.” *Id.*, [0040], [0043]-[0044]; *see also* ’374 patent, 5:25-27 (“Thus, the light-emission control unit 105 turns off the light-emitting device 106 with no drive current supplied thereto.”). In particular, IBM discloses operating in “detection mode”—i.e., standby mode—wherein the microphones deactivate and transmission of audio data ceases for recording, speech processing, and providing feedback ceases until speech is detected. IBM, [0043]-[0044]. Thus, when operating in detection mode, IBM’s system does not provide feedback to users using the lighted bar such that it is “turned off.” IBM, [0040], [0043]-[0044]. This aligns with the ’374 patent’s supporting description. ’374 patent, 5:25-27 (“Thus, the light-emission control unit 105 turns off the light-emitting device 106 with no drive current supplied thereto.”). IBM’s system remains in detection mode until detecting the onset of speech activity based on, e.g., “user activation of a microphone control button” or sensed conditions (e.g., visual or biometric sensing). *Id.*, [0043]-[0046], [0020] (“[S]peech activity events include user interactions with microphone switch buttons, speech intent

events that indicate a user's intention to speak into the microphone (as determined based on processing of biometric feature data of the user and/or from data acquired from tracking behavioral characteristics of the user, etc.”)

(2) [7b] the light-emitting device is turned on if the communication mode is the transmission mode and if the speech signal to be transmitted is the speech segment, and if the speech quality of the speech signal to be transmitted is evaluated as bad, and the light-emitting device is repeatedly turned on and off if the communication mode is the transmission mode and if the speech signal to be transmitted is the speech segment, and if the speech quality of the speech signal to be transmitted is evaluated as good.

[0130] IBM renders [7b] obvious because it discloses varying the brightness of its lighted bar to track voice activity and indicating voice quality using different colors (as I also discuss with respect to [1c]). IBM uses its feedback generation system (114) in transmission mode to control audio quality feedback displayed to a user on its lighted bar. IBM, [0012]; [0023], [0027], [0036], [0039], and [0040]. In one example, IBM discloses turning feedback output device “on” to indicate the SNR or other factor(s) is impacting the system’s ability to identify speech segments or obtain clear speech sounds. IBM, [0040]. IBM also explains that “the system can provide an indication to the user that the system is listening and capturing the speech

input when a length of brightness of a lighted bar quickly tracks the user's voice activity." IBM, [0040]. Thus, the lighted bar also "repeatedly turns off and on" as it tracks the user's voice, turning "off" between speech segments and back "on" again as the system recognizes and evaluates spoken words.

[0131] In one example, IBM discloses turning red or yellow lights of the feedback output device "on" to indicate the SNR and/or other factors are impacting the system's ability to identify speech segments or obtain clear speech sounds. IBM, [0040]. IBM also explains that when it can distinguish speech segments, "the system can provide an indication to the user that the system is listening and capturing the speech input when a length of brightness of a lighted bar quickly tracks the user's voice activity." Id., [0040]. Thus, the lighted bar also "repeatedly turns off and on" as it tracks the user's voice, turning "off" between speech segments and back "on" again as the system recognizes and evaluates spoken words. Indeed, IBM's lighted bar provides feedback for each of the '374 patent's described LED states corresponding to sound pick-up state determinations:

	COMMUNICATION MODE	SOUND PICK-UP STATE DETERMINATION		LED STATE
		SPEECH SEGMENT DETERMINATION	SPEECH QUALITY DETERMINATION	
No1	STANDBY MODE IBM: Detection Mode	NO DETERMINATION	NO DETERMINATION	TURN OFF
No2	TRANSMISSION MODE	NON-SPEECH SEGMENT	NO DETERMINATION	TURN ON RED
No3		SPEECH SEGMENT	BAD SPEECH QUALITY	TURN ON YELLOW
No4			GOOD SPEECH QUALITY	TURN ON AND OFF GREEN

'374 Patent, Fig. 13 (Annotated with IBM Lighted Bar)

[0132] Moreover, IBM further discloses turning a lighted bar “on” to notify the user of voice volume quality determinations using different colors, e.g., green, yellow, or red to indicate a normal, borderline, or overly loud signal. Moreover, as discussed regarding [1c], one of skill would also recognize that the claimed on-off cannot distinguish the claims from the prior art. *See, e.g.*, Section V. This also aligns with the '374 patent's disclosure. '374 patent, 39:48-40:2. Thus, one of ordinary skill would also understand that IBM discloses [7b] by turning the lighted bar “on” and “off” when in the transmission mode to notify the user of speech quality by controlling the brightness of the bar and/or providing different colors.

5. Claims 6 and 12

a) Claim 6

(1) [6a] The condition notification method according to claim 1 further comprising a microphone-direction determination step of determining a direction of a main microphone based on a phase difference of sounds picked up by a first sound pick-up unit and a second sound pick-up unit both provided for the simplex communication apparatus,

[0133] In my opinion, the combination of IBM and Kale renders [6a] obvious because Kale detects the arrival direction of speech based on the “spatial orientation” of received sound signals. As I discussed for element [4a], both IBM and Kale improve acoustic quality by cancelling noise from a speech signal received at one microphone based on noise detected by another microphone. *See* IBM, [0037], [0053]; Kale, Abstract, [0003], [0028] [0043], FIGS. 8-9; *see also* ’374 patent, 1:29-32 (Background of the Invention).

[0134] Kale’s noise cancelling device further includes a “direction of arrival (DOA) module” for “determining the direction from which the speech data arrived.” Kale, [0028], [0031], [0046], FIGS. 8-9. The DOA module detects the arrival direction of the speech by “comput[ing] a cross-correlation between the two [sound] signals 845, 850” that are based on the raw audio data and determining a “spatial orientation” of the signals. Kale,

[0003], [0031].

[0135] One of ordinary skill would have understood that the computed “cross correlation” pertains to a phase-difference calculation, i.e., the difference between the angle at which the first and second sound signals are received with respect to a reference microphone. In fact, Kale explains that the DOA module’s output includes the phases of the received sounds signals, i.e., “two angles in degrees, the first being for a desired speech signal while the second is for noise.” Kale, [0031]. Specifically, Output “ θN is the direction of noise and [output] θS is the direction of sound.” Kale, [0039].

[0136] Kale thus discloses that its DOA module detects the voice’s incoming direction to the main microphone based on a phase difference between the first and second sound signals from its dual microphone setup. And, in my opinion, one of ordinary skill would have understood that these angles described by Kale correspond to the phases of the first and second sound signals—that is, θN corresponds to the second sound signal when it contains the noise signal and θS corresponds to the first sound signal when it contains the voice signal.

[0137] One of ordinary skill would have therefore implemented Kale’s noise-suppression techniques based on voice-direction to supplement

IBM's existing noise reduction techniques and further enhance voice sound quality. Indeed, a POSITA would have been motivated to consider their combined teachings given IBM and Kale disclose similar systems that operate in a similar manner and are directed to similar problems.

(2) [6b] the first sound pick-up unit including the main microphone, wherein, in the speech-quality evaluation step, the speech quality of the speech signal to be transmitted is evaluated based on the direction of the main microphone thus determined in the microphone-direction determination step.

[0138] It is also my opinion that IBM combined with Kale discloses or renders obvious [6b]. IBM's system detects the existence of one or more conditions leading to degraded acoustic quality of the captured audio. IBM, [0047]. Like Kale, IBM discloses a microphone system (101) with "a microphone that is dedicated to capturing environmental noise and another microphone for capturing speech input." IBM, [0053]. "For instance, the environment interpretation system (113) can determine or otherwise detect the static environmental noise in an audio signal received via a user microphone or a microphone specifically implemented as a detector of the non-speech environment and generate and update models/parameters of the environmental noise" IBM, [0037], [0038], [0043], [0053]. Arguably, however, IBM does not explicitly disclose evaluating speech based on the

direction of the main microphone.

[0139] IBM does disclose, however, that its “acoustic quality control functions are supported by system modules,” IBM, [0036], aligning with Kale’s explanation that its noise cancelling device “may be incorporated with any electronic device that includes an audio receiving device such as a microphone.” Kale, [0028]. Kale’s noise canceling device also includes a “beamformer” and “IFFT module” for “modify[ing] an audio signal to isolate a speech therein” or “enhance[ing] a speech of an audio signal” based on voice direction. Kale, [0034]. In particular, Kale explains that the beamformer/ IFFT module reduces the noise component of whichever signal contains the speech signal (either the first or second sound signal) using the delayed noise signal (the other of the first or second sound pick up signal). Kale, [0047]-[0048], FIGS. 8 and 9. Accordingly, one of ordinary skill would have been motivated to implement Kale’s noise identification and reduction techniques based on direction of the main microphone as part of its “acoustic quality control functions are supported by system modules.” IBM, [0036]-[0038]; Kale, [0028], [0047]-[0048], FIGS. 8 and 9.

b) Claim 12

(1)[12a] The communication apparatus according to claim 11 further comprising a microphone-direction determination unit configured to determine a direction of a main microphone based on a phase difference of the voice sounds picked up the first and second sound pick-up units, respectively,

[0140] Kale renders [12a] obvious because it discloses a direction of arrival (DOA) module, as I discuss in further detail with respect to [6a]’s “a microphone-direction determination step of determining a direction of a main microphone based on a phase difference of sounds picked up by a first sound pick-up unit and a second sound pick-up unit both provided for the simplex communication apparatus.”

(2)[12b] the first sound pick-up unit including the main microphone, wherein, the speech-quality evaluation unit evaluates the speech quality of the speech signal to be transmitted by the transmitter unit based on the direction of the main microphone thus determined by the microphone-direction determination unit.

[0141] Kale renders [12b] obvious, as I discuss in further deal [6b]’s “the first sound pick-up unit including the main microphone, wherein, in the speech-quality evaluation step, the speech quality of the speech signal to be transmitted is evaluated based on the direction of the main microphone thus determined in the microphone-direction determination step.”

6. Claims 8 and 15

a) Claim 8

(1)(a) [8a] The condition notification method according to claim 1, wherein the speech-segment determination step includes: a frequency conversion step of converting a signal input from a sound pick-up unit provided for the simplex communication apparatus into a signal component in unit of a specific length in a frequency domain; and

[0142] In my opinion, IBM and Kale alone or in combination disclose, or at least suggest, [8a] because both reference speech recognition techniques in the frequency domain to identify speech segments. IBM, [0026]; [0029]; Kale, [0003], [0043]-[0044]; *see also* Technology Background; Exs. 1009-1010. IBM explains that ASR system 120 provides support for speech-enabled applications using known automatic speech recognition (ASR) techniques. IBM, [0026]. One of ordinary skill would appreciate from IBM's disclosure, particularly its discussion of ASR 120, speech energy, etc., that existing spectral analysis techniques were both understood and readily available for evaluating captured audio signals to identify and distinguish voice and speech characteristics. IBM, [0026]; [0029] ("The audio data processing system (116) may implement known audio-based techniques for processing the audio input," including determinations regarding speech energy.)

[0143] Accordingly, one of ordinary skill would have understood IBM's disclosure encompasses well-known Fast Fourier Transform (FFT) algorithms for converting audio signals to their frequency domain equivalent to identify and distinguish speech. Among other things, one of skill would have understood transitioning to the frequency domain allows for the detection of speech onset and offset by analyzing the voice signal for frequency components of a specific length.

[0144] To remove all doubt, Kale provides explicit disclosure. As Kale explains, it distinguishes voice from noise in the frequency domain to identify speech:

The FFT 805 may receive a first microphone speech data 835 while the FFT810 may receive a second microphone speech data 840. With reference to the exemplary rate of 20 ms, speech samples from the first and second microphones in 20 ms frames are computed by the FFTs 805, 810, respectively. According to the exemplary embodiments, the FFTs 805,810 may compute a 128, 256, and/or 512 point FFT of a 8 kHz signal, thereby breaking into 64, 128, and/or 256 frequency bins. Again, it should be noted that the computations of the FFTs 805, 810 is only exemplary and the computations may be changed as a function on the resolution desired and the platform capabilities to handle the FFTs processing. For example, if a 128 point FFT is selected, 64 frequency bins from 0-4000 Hz are generated.

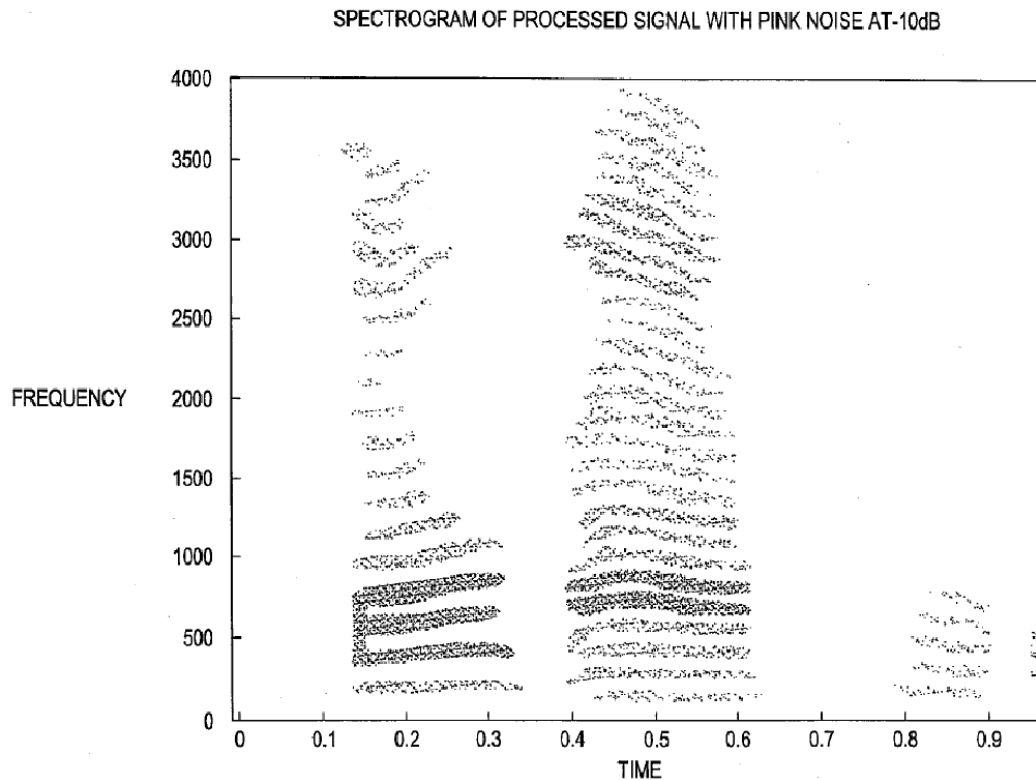
Kale, [0029]. The computation module, which may be an FFT module, “determine[s] which of the raw audio data signals comprises the speech signal” using “factors ... such as common formants, formants with known patterns, etc.” Kale, [0003], [0043]-[0044]]; *see also* [0003] (“A first transformed signal relates to speech data”), claim 1 (“first computation modules receiv[e] raw audio data and generat[e] a respective transformed signal as a function of formants, a first transformed signal relating to speech data), claim 3 (“wherein the first computation modules are Fast Fourier Transform (FFT) modules.”). In my opinion, one of ordinary skill would consider these teachings of Kale with IBM to “eliminate a maximum amount of the noise data by performing noise cancellation “as a function of formant shaping and prioritizing cancellation of noise at frequencies that are more sensitive over noise at frequencies that are less sensitive to noise is desired, thereby leading to significantly improved audio performance.” Kale, [003], [022].

(2) [8b] a sound determination step of determining whether the signal input from the sound pick-up unit carries a voice component or a noise component based on a spectrum component of the signal component thus converted into the frequency domain.

[0145] IBM and Kale also disclose, or at least suggest, [8b] because both recognize speech recognition techniques in the frequency domain to

distinguish speech and noise components. IBM discloses a microphone system having “a microphone that is dedicated to capturing environmental noise and another microphone for capturing speech input.” IBM, [0053]. IBM also explains that “the environment interpretation system (113) can determine or otherwise detect the static environmental noise in an audio signal received via a user microphone or a microphone specifically implemented as a detector of the non-speech environment and generate and update models/parameters of the environmental noise (117c)” IBM, [0037], [0038], [0043], [0053].

[0146] Kale further discloses details regarding operations in the frequency domain. Kale, [0025]. Kale further discloses details known to those of skill in the art regarding operations for operating in the frequency domain. Kale, [0025]. For example, Kale explains that even though the spatial orientations of beams of microphones can at least partially reduce noise, this “does not account for the psychoacoustics fact that the spatial intensity direction and frequency intensity direction for noise is not always connected.” Kale, [022]. Kale goes on to describe a spectrogram of speech in Figure 6 involving an “interfering talker and pink noise coming from a single location in space.” Kale, [0025].



Kale, Fig. 6.

[0147] One of ordinary skill would readily appreciate from Figure 6 “between 0.2-0.3 seconds, the maximum intensity is around 500 Hz while between 0.4-0.5 seconds, the intensity is around 500 Hz as well as 2000 HZ and 3000 HZ,” reflecting that “intensity is different at different frequencies and changes with time.” Kale, [022]. As taught by Kale, one of ordinary skill would consider these teachings to “eliminate a maximum amount of the noise data by performing noise cancellation “as a function of formant shaping and prioritizing cancellation of noise at frequencies that are more sensitive over noise at frequencies that are less sensitive to noise is desired,

thereby leading to significantly improved audio performance.” Kale, [003], [022]. Thus, in my opinion, Kale establishes one of ordinary skill art understood how to use speech recognition techniques in the frequency domain to distinguish speech and noise components. In my opinion, one of ordinary skill would thus understand IBM and Kale to each disclose, or at least suggest, element [8b].

b) Claim 15

(1)[15a] The communication apparatus according to claim 9, wherein the speech-segment determination unit converts the first signal input from the first sound pick-up unit into a signal component in unit of a specific length in a frequency domain and

[0148] IBM and Kale each disclose, or at least suggest, [15a]

because both refer to using speech recognition techniques in the frequency domain for identifying speech segments, as I discuss as it relates to [8a].

(2)[15b] determines whether the first signal carries a voice component or a noise component based on a spectrum component of the signal component thus converted into the frequency domain.

[0149] IBM and Kale each disclose, or at least suggest, [15b]

because both recognize speech recognition techniques in the frequency domain to distinguish speech and noise components, as I discuss as it relates to [8a].).

C. Ground 3: IBM, Kale, and Li Render Claim 14 Obvious**1. Overview of Li**

[0150] I understand that because Li was filed July 28, 2009, and published February 3, 2011, qualifies as prior art to the '374 patent under §§102(b). Li discloses an “audio processing apparatus” for “suppressing stationary and non-stationary noises.” Li, [0019]. The apparatus includes two microphones (a main and a reference microphone), a voice active detector, a beamformer, and a noise suppressor. Li, [0019], FIG. 2A. Noting the importance of the microphones’ relative location for signal processing purposes, Li explains that its main and reference microphone are disposed at different locations such that the main microphone is closest to the speaker. Li, [0020]. Li also frames its noise-suppression discussion with respect to a cell phone, explaining that “the cellphone 300 comprises a front panel 310 and a back panel 320 and the main microphone 202 is disposed on the bottom of the front panel 310, and the reference microphone 204 is disposed on the top of the back panel 320.” Li., [0020].

[0151] Li’s main and reference microphones receive a “mixture of sounds”—from a “source” (such as a “speaker’s mouth”) and a “non-source” (e.g. “noises”) before “convert[ing] the mixture of the sounds and the noises [] into a main input M1 and a reference input M2.” Li, [0020]. The

converted inputs are then converted into analog signals (S_1 , S_2), calibrated (into signals C_1 and C_2), and input into the voice active detector (VAD), which uses the calibrated signals DOA estimation “to generate a voice active signal V_1 .” Li, [0020], [0025], FIG. 2A.

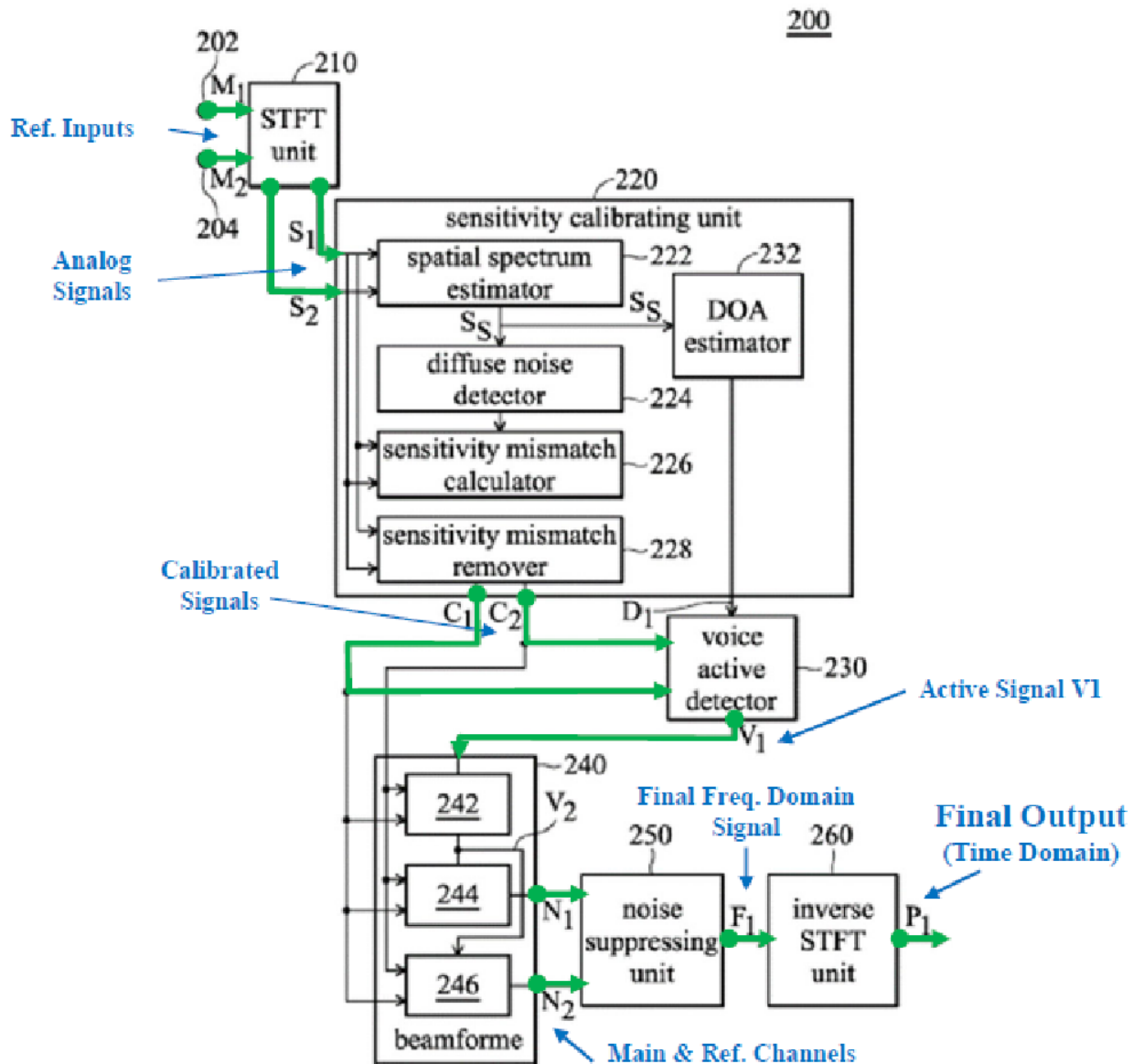


FIG. 2A

Li, FIG. 2A.(Annotated)

[0152] Li's beamformer uses the voice active signal V1 to generate and output "steering vector signal V2," which "indicates [the] directional vector" of the source signal (voice signal) when voice is detected. Li, [0026], [0031], claim 13, FIGS. 2A, 3C. The beamformer also generates "main" and "reference" channels N1 and N2 ("corresponding to the sounds received from the source" and "the noises received from non-source," respectively) and outputs the signals to the noise suppressing unit. Li, [0026]-[0027], FIG. 2A, 3D. Finally, the noise suppressing unit "suppress[es] stationary and non-stationary noises in the main channel N1 and the reference channel N2 according to the voice active signal V1, and integrate[s] the main channel N1 and the reference channel N2 into a final signal F1." Li, [0027], FIGS. 2A, 3A. Li then applies an inverse Short Time Fourier Transformation (STFT) to "convert the final signal F1 of the frequency domain signals into a final output P1 of the time." Li, [0027], FIGS. 2A, 3A.

2. Reasons to Combine

[0153] In my opinion, IBM, Kale, and Li are analogous art to each other and the '374 patent because they are in the same field of endeavor: audio signal processing. Li, [0020] ("The placement of the two microphones is advantageous for signal processing"); IBM, [0002]-[0008], [0025]; '374 patent, 1:17-2:35, 38: 3-13; Kale, Abstract. IBM and Li also share another

common aim, namely to enhance voice sound quality by automatically calibrating microphone sensitivity to reduce noise in a double microphone setup. IBM, [0006], [0013]; Li, [0002], [0007]-[0008], [0022]-[0024]. Li's disclosure merely provides specific locations for the microphones when using these additional techniques based on the direction of the speaker relative to the main microphone.

[0154] For example, IBM explains that already traditional ASR systems “determine[] the volume of the speaker's voice relative to the static environmental noise, which is then used to adjust the sensitivity and volume of the microphone input.” IBM, [0006]. IBM goes on to explain that its “the acoustic quality control system can detect a change in an environmental static noise level [based on a determined distance from the speaker] and automatically adjust a microphone volume sensitivity parameter.” IBM, [0013], [0037], [0054]-[0055].

[0155] Likewise, Li explains that its disclosure “relates to an audio processing apparatus and method for microphone sensitivity calibration.” Li, [0020]. And like IBM and the '374 patent, Li provides a main microphone for voice and another for “noises from non-source sources. *Compare* Li, [0020] with '374 patent, 1:29-51, 7:37-44 and IBM, [0038], [0046], [0053], Fig. 2. And with this double microphone setup, “[t]he present invention uses

the diffuse noises to calibrate the sensitivity mismatch between the microphones 202 and 204.” Li, [0024]. However, Li adjusts its microphone sensitivity based on the direction from which the sound came.

[0156] At least because IBM, Kale, and Li disclose similar systems that operate in a similar manner and are directed to similar problems as the ’374 patent, it is my opinion that a POSITA would have been motivated to consider their combined teachings. *Compare, e.g.*, Li, [0020] and IBM, [0002]-[0008], [0025] *with* ’374 patent, 1:17-2:35, 38:3-13.

3. Claim 14: The communication apparatus according to claim 11, wherein the first and second sound pick-up units are a main microphone and a sub-microphone, respectively, and the main microphone and the sub-microphone are arranged at a front face and a rear face of the communication apparatus, respectively.

[0157] Like IBM, Kale, and the ’374 patent, Li describes a noise cancellation system having a double microphone setup. Li, [0021]. As the ’374 patent recognizes, the claimed dual microphone setup was already well-known at the time of filing. ’374 patent, 1:29-32. Moreover, the ’374 patent’s “BACKGROUND OF THE INVENTION” further confirms a POSITA already appreciated the claimed techniques for noise-cancellation based on the positional relationship between the user’s mouth and a “handset microphone,” also referred to as a telephone receiver. *Id.*, 1:33-42.

[0158] Regardless, and as illustrated in Figure 2B below, Li’s

apparatus includes a “main microphone” and a “reference microphone” (first and second microphones) disposed along front and back panels of, e.g., a cell phone. Li, [0020], FIGS. 2A, 2B.

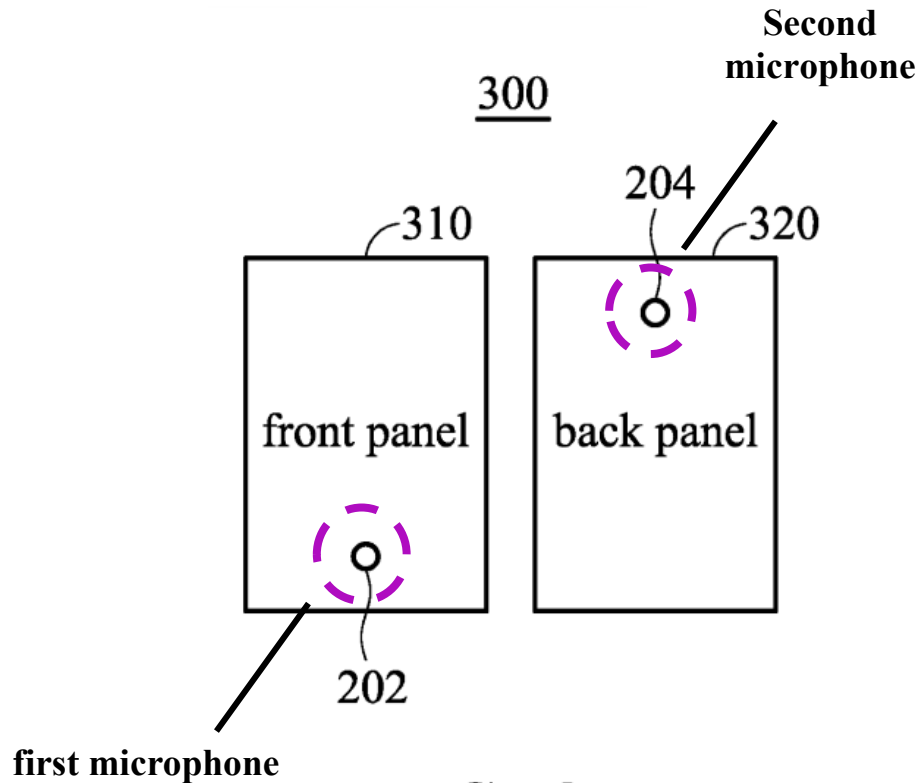


FIG. 2B

Li, Fig. 2B (Annotated with First and Second Microphones)

[0159] Li thus discloses main microphone and a sub-microphone arranged at a front face and a rear face of a communication apparatus, respectively. One of ordinary skill would have been motivated to modify the IBM-Kale combination to follow Li’s teaching on microphone placement to better distinguish speech from noise signals. For example, one of skill would

have appreciated Li's explicit reasoning for locating a microphone at the front and back face of the device:

The main microphone 202 is closer to the source, e.g. a speaker's mouth, than the reference microphone 204. Note that there is also a physical barrier between the front side 310 and the back side 320 so the reference microphone 204 may capture less Sounds from the source than the main microphones 202. The placement of the two microphones is advantageous for signal processing. In this embodiment, the main microphone 202 and the reference microphone 204 will respectively convert the mixture of the Sounds and the noises *t* into a main input M1 and a reference input M2 as shown in FIG. 2.

Li, [0020].

[0160] IBM, Kale, and Li use similar element and techniques to achieve similar results, and it is my opinion a POSITA would have had a reasonable expectation of success in configuring the IBM-Kale combination to implement Li's microphone placement on the front and back of the device.

IX. Conclusion

[0161] This declaration and my opinions herein are made to the best of my knowledge and understanding, and based on the material available to me, at the time of signing this declaration. I declare that all statements made herein on my own knowledge are true and that all statements made on

information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 or Title 18 of the United States Code.

Date: February 7, 2025

Stuart J. Lipoff
Stuart Lipoff