

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

HARMAN INTERNATIONAL INDUSTRIES, INC.,

Petitioner,

v.

ST CASESTECH LLC,

Patent Owner.

Case IPR2024-01303

U.S. Patent No. 11,521,632

**DECLARATION OF DR. MICHAEL T. JOHNSON IN SUPPORT
OF PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 11,521,632**

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. QUALIFICATIONS.....	4
III. LEGAL UNDERSTANDING.....	8
IV. BACKGROUND OF THE TECHNOLOGY.....	14
V. THE '632 PATENT	16
A. '632 Patent Overview.....	16
B. The Prosecution History.....	17
I. LEVEL OF ORDINARY SKILL IN THE ART	19
II. CLAIM CONSTRUCTION.....	20
VI. GROUNDS OF UNPATENTABILITY.....	20
VII. SPECIFIC GROUNDS FOR CHALLENGE	22
A. GROUND 1: Victorian in view of Jannard (Claims 1-7, 9, 15-16)....	23
1. Overview of Victorian.....	23
2. Overview of Jannard	26
3. Motivation to Apply Jannard's Teachings to Victorian	29
4. Invalidity of Claims 1-7, 9, and 15-16 Over Victorian in view of Jannard.....	37
B. GROUND 2: Ground 1 in further view of Karaoguz (Claim 8):	96
1. Overview of Karaoguz	96
2. Motivation to Apply the Karaoguz's Teachings to Victorian and Jannard	97
3. Claim [8].....	99
C. GROUND 3: Ground 1 in further view of Nathan (Claims 10-14):	102

1.	Overview of Nathan	102
2.	Motivation to Apply Barnes’ Teachings to Jannard and Wittenberg.....	102
3.	Claims [10] and [12].....	Error! Bookmark not defined.
4.	Claim [11].....	104
5.	Claims [13] and [14].....	109
D.	GROUND 4: Ground 1 in further view of Camp (Claims 16-17):...	110
1.	Overview of Camp	110
2.	Motivation to Apply the Camp’s Teachings to Victorian and Jannard	111
3.	Claim [16].....	113
4.	Claim [17].....	114
VIII.	SECONDARY CONSIDERATIONS	117
IX.	CONCLUSION	117

I. INTRODUCTION

1. My name is Dr. Michael T. Johnson and I am a Full Professor in the Department of Electrical and Computer Engineering at the University of Kentucky in Lexington, Kentucky. I have been retained on behalf of Harman International Industries, Incorporated (“Harman”), to provide my opinions on certain issues related to U.S. Patent No. 11,521,632 (“the ’632 patent”) in connection with the above-captioned inter partes review (IPR) proceeding. In particular, I have been asked to provide my insights, analysis, and opinions regarding whether claims 1-17 of the ’632 patent (“Challenged Claims”) are obvious over the prior art references identified below.

2. I understand that the ’632 patent is titled “Personal Audio Assistant Device and Method,” names as inventor Steven Wayne Goldstein, and is currently owned by ST CasesTech LLC. I have considered the ’632 patent. I understand that the ’632 patent has been provided as Ex. 1001.

3. I understand that the file history of the ’632 patent has been provided as Ex. 1002. I have considered this file history, and I will refer to it as “the ’632 File History.”

4. I understand that the ’632 patent is a continuation of U.S. Patent Application No. 14/148,749, which is a continuation of U.S. Patent Application No. 11/774,965, which claims priority to provisional application No. 60/806,769. I

understand that the file history for the provisional applications No. 60/806,769 has been provided as Exhibit 1005. I have considered this file history.

5. I understand that the '632 patent claims a priority date of July 8, 2006, the filing date of provisional application No. 60/806,769. I have been asked to assume that priority date for purposes of rendering my opinions in this declaration. I have not been asked to opine on whether the claims of the '632 patent are entitled to this priority date, and do not provide any opinion on that subject.

6. My analysis and conclusions set forth in this declaration are based on my educational background and experiences in the field (see Section II). I am qualified to opine concerning what a person of ordinary skill in the art would have known and understood during that time. Before July 8, 2006, my level of skill in the art was at least that of a person of ordinary skill. I am qualified to opine concerning what a person of ordinary skill in the art would have known and understood at the time of the '632 patent's purported invention. My analysis and conclusions herein are from the perspective of a person of ordinary skill in the art as of July 8, 2006.

7. As part of my independent analysis for this declaration, I have considered the following: the background knowledge/technologies that were commonly known to persons of ordinary skill in this art during the time before the earliest claimed priority date for the '632 patent; my own knowledge and

experiences gained from my work experience in the field of the '632 patent and related disciplines; and my experience in working with others involved in this field and related disciplines.

8. I have considered the following non-exhaustive list of references in preparing my declaration. I have been instructed to assume that each of these references qualifies as prior art.

- **European Publication No. EP1519625** to Thomas A. Victorian (“Victorian”), titled “External ear canal voice detection” filed on September 13, 2004, and published in March 30, 2005. I understand that a copy of this reference has been designated as Exhibit 1007.
- **U.S. Patent Application Publication No. 2006/0132382** to James H. Jannard (“Jannard”), titled “Data input management system for wearable electronically enabled interface” filed on Dec. 22, 2004, and published Jun. 22, 2006. I understand that a copy of this reference has been designated as Exhibit 1008.
- **U.S. Patent Application Publication No. 2004/0117846** to Jeyhan Karaoguz (“Karaoguz”), titled “Personal access and control of media peripherals on a media exchange network” filed on September 11, 2003, and published June 17, 2004. I understand that a copy of this reference has been designated as Exhibit 1009.
- **U.S. Patent Application Publication No. 2004/0167847** to Alan B. Nathan (“Nathan”), titled “System and method for trading financial instruments using speech” filed on August 15, 2003, and published August 26, 2004. I understand that a copy of this reference has been designated as Exhibit 1010.
- **U.S. Patent Application Publication No. 2007/0274531** to William O. Camp (“Camp”), titled “Sound pressure monitor” filed on May 24, 2006, and published November 29, 2007. I understand that a copy of this reference has been designated as Exhibit 1011.

9. I understand that there is a pending district court litigation between

Petitioner and ST CasesTech, LLC (“Harman Litigation”) involving the ’632 patent.

10. I am being compensated by Harman at my standard hourly consulting rate of \$750 for my time on this matter. My compensation is not dependent on the outcome of this proceeding or any other proceeding between the parties.

11. In forming my opinions, I relied on the documents cited in this declaration and the documents identified in the attached Appendix B. These documents comprise patents, file histories, printed publications, and other related documents. As discussed below, each document is a type that experts in my field would have reasonably relied upon when forming their opinions. Further, experts in my field would have had access to each document either through the applicable patent offices and/or well-known libraries, conferences, or publications in the field. My opinions are also based upon my personal and professional experience.

II. QUALIFICATIONS

12. I am qualified by education and experience to testify as an expert in the field of audio devices and audio signal processing. Additionally, the following overview of my background pertains to my qualifications for providing expert testimony in this matter. I am over the age of 18 and am competent to write this declaration. I have personal knowledge, or have developed knowledge, of these

technologies based upon education, training, or experience of the matters set forth herein.

13. My educational background includes undergraduate degrees both in Computer Science Engineering and in Electrical Engineering from LeTourneau University in Longview, TX, in 1989 and 1990, respectively, and a Master's degree in Electrical Engineering from the University of Texas at San Antonio in 1994. I earned a PhD from Purdue University in West Lafayette, Indiana, in 2000, with a doctoral dissertation focused on speech processing and automatic speech recognition. Prior to receiving my PhD, I worked in industry as a design engineer and engineering manager for more than 5 years, primarily in the telecommunications industry.

14. I am currently a Full Professor in the Department of Electrical and Computer Engineering at the University of Kentucky (UK) in Lexington, Kentucky. I also currently serve as the Associate Dean for Undergraduate Education and Student Success in the Stanley and Karen Pigman College of Engineering at the UK, a role I have held since July 2023.

15. I previously served as the Electrical and Computer Engineering Department Chair from 2016 to 2023.

16. Prior to my faculty role at UK, I was a faculty member at Marquette University in Milwaukee, WI from 2000 to 2016, during which time I also spent two

separate sabbatical years as a Visiting Professor at Tsinghua University in Beijing, China.

17. The focus of my research work is Speech and Signal Processing, and I regularly conduct and publish original research articles within this field, with a publication record including over 50 peer-reviewed journal articles and more than 150 total peer-reviewed publications over the past three decades. My specific expertise includes automatic speech recognition, speech production and kinematics, bioacoustics, and machine learning. I have received multiple research grant awards for my work in this area, including from the National Science Foundation and the National Institutes of Health.

18. As a faculty member, I regularly teach both undergraduate and graduate courses in electrical and computer engineering, primarily in the areas of computers, signal processing, and machine learning. This includes graduate level courses in speech and signal processing, including topics directly related to acoustics and acoustic devices such as those at issue in these proceedings.

19. My professional service includes active membership in the Institute for Electrical and Electronic Engineers (IEEE), the IEEE Signal Processing Society, the Acoustical Society of America (ASA), and the Electrical and Computer Engineering Department Heads Association (ECEDHA). I served as a member of the IEEE Signal Processing Society Speech and Language Technical Committee from 2014 to

2017, and regularly serve as a reviewer for journals, conferences, and workshops in the area of speech and signal processing, including IEEE Transactions on Acoustics, Speech, and Signal Processing, IEEE Transactions on Signal Processing, the Journal of the Acoustical Society of America, Speech Communication, the International Conference on Acoustics, Speech, and Signal Processing, the IEEE Workshop on Automatic Speech Recognition, and Interspeech. I have reviewed funding proposals for a number of public and private agencies and foundations, including for multiple programs and directorates at the National Science Foundation. In addition to these manuscript and grant proposal reviewing activities, I also currently serve as Area Editor for the journal Speech Communication in the area of speech enhancement.

20. During my academic career, I have been honored to receive a number of awards for my teaching and research work. Notable awards have included the College of Engineering Researcher of the Year at Marquette University in 2007, Milwaukee Young Engineer of the Year in 2006, and Electrical and Computer Engineering Eta Kappa Nu Teacher of the Year in 2005 and 2014. In 2008, my work applying speech recognition technology to animal vocalizations was featured in the Milwaukee Journal-Sentinel series “Groundbreaking Thinkers in Wisconsin.”

21. A detailed record of my professional qualifications, including a list of publications, awards, and professional activities, is set forth in my curriculum vitae attached to this report as Appendix A.

III. LEGAL UNDERSTANDING

A. My Understanding of Claim Construction

22. I am not an attorney and offer no legal opinions, but in the course of my work, I have had experience studying and analyzing patents and patent claims from the perspective of a person skilled in the art.

23. I have been informed that patent claims are construed from the viewpoint of a person of ordinary skill in the art of the patent at the time of the invention. I have been informed that patent claims generally should be interpreted consistent with their plain and ordinary meaning as understood by a person of ordinary skill in the art in the relevant time period (i.e., at the time of the purported invention, or the so called “effective filing date” of the patent application), after reviewing the patent claim language, the specification, and the prosecution history (i.e., the intrinsic record).

24. I have further been informed that a person of ordinary skill in the art must read the claim terms in the context of the claim itself, as well as in the context of the entire patent specification. I understand that in the specification and prosecution history, the patentee may specifically define a claim term in a way that differs from the plain and ordinary meaning. I understand that the prosecution history of the patent is a record of the proceedings before the U.S. Patent and Trademark Office and may contain explicit representations or definitions made

during prosecution that affect the scope of the patent claims. I understand that an applicant may, during the course of prosecuting the patent application, limit the scope of the claims to overcome prior art or to overcome an examiner's rejection, by clearly and unambiguously arguing to overcome or distinguish a prior art reference, or clearly and unambiguously disavowing claim coverage.

25. In interpreting the meaning of the claim language, I understand that a person of ordinary skill in the art may also consider "extrinsic" evidence, including expert testimony, inventor testimony, dictionaries, technical treatises, other patents, and scholarly publications. I understand this evidence is considered to ensure that a claim is construed in a way that is consistent with the understanding of those of ordinary skill in the art at the time of the claimed invention. For example, this can be useful for a technical term whose meaning may differ from its ordinary English meaning. I understand that extrinsic evidence may not be relied on if it contradicts or varies the meaning of claim language provided by the intrinsic evidence, particularly if the applicant has explicitly defined a term in the intrinsic record.

B. My Understanding of Anticipation

26. I understand that a patent claim is invalid as "anticipated" if all the limitations of the claim are present, either expressly or inherently, in a single prior art publication or patent.

27. I also understand that a prior art reference can disclose a claimed feature because the feature is expressly described or because the feature is inherent in the disclosure. I understand that something is inherent in a prior art reference if the missing descriptive matter must necessarily be present, not merely probably or possibly present, and it would be so recognized by a person of ordinary skill in the art as of the priority date of the patent. I understand that it is acceptable to examine evidence outside of the prior art reference (i.e., extrinsic evidence) in determining whether a feature, while not expressly discussed in the reference, is necessarily present in it.

C. My Understanding of Obviousness

28. I understand that a claim may be invalid if the subject matter described by the claim as a whole would have been obvious to a hypothetical person of ordinary skill in the art in view of a prior art reference or in view of a combination of references at the time the claimed invention was made. Therefore, I understand that obviousness is determined from the perspective of a hypothetical person of ordinary skill in the art and that the asserted claims of the patent should be read from the point of view of such a person at the time the claimed invention was made. I further understand that a hypothetical person of ordinary skill in the art is assumed to know and to have all relevant prior art in the field of endeavor covered by the patent in suit.

29. I have been informed that there are two criteria for determining whether prior art is analogous and thus can be considered prior art: (1) whether the art is from the same field of endeavor, regardless of the problem addressed, and (2) if the reference is not within the field of the patentee's endeavor, whether the reference still is reasonably pertinent to the particular problem with which the patentee is involved. I have also been informed that the field of endeavor of a patent is not limited to the specific point of novelty, the narrowest possible conception of the field, or the particular focus within a given field. I have also been informed that a reference is reasonably pertinent if, even though it may be in a different field from that of the patentee's endeavor, it is one which, because of the matter with which it deals, logically would have commended itself to a patentee's attention in considering their problem.

30. I have also been advised that an analysis of whether a claimed invention would have been obvious should be considered in light of the scope and content of the prior art, the differences (if any) between the prior art and the claimed invention, and the level of ordinary skill in the pertinent art involved. I understand as well that a prior art reference should be viewed as a whole.

31. I have also been advised that in considering whether a claimed invention would have been obvious over a combination of prior art references, I may assess whether there are apparent reasons to combine known elements in the prior

art in the manner claimed in view of interrelated teachings of multiple prior art references, the effects of demands known to the design community or present in the marketplace, and/or the background knowledge possessed by a person of ordinary skill in the art. I understand that other principles may be relied on in evaluating whether a claimed invention would have been obvious, and that these principles include the following:

- A combination of familiar elements according to known methods is likely to be obvious when it does no more than yield predictable results;
- When a device or technology is available in one field of endeavor, design incentives and other market forces can prompt variations of it, either in the same field or in a different one, so that if a person of ordinary skill can implement a predictable variation, the variation is likely obvious;
- If a technique has been used to improve one device, and a person of ordinary skill in the art would recognize that it would improve similar devices in the same way, using the technique is obvious unless its actual application is beyond their skill;
- An explicit or implicit teaching, suggestion, or motivation to combine two prior art references to form the claimed combination may demonstrate obviousness, but proof of obviousness does not depend on or require showing a teaching, suggestion, or motivation to combine;
- Market demand, rather than scientific literature, can drive design trends and may show obviousness;
- In determining whether the subject matter of a patent claim would have been obvious, neither the particular motivation nor the avowed purpose of the named inventor controls whether the claim is obvious;
- One of the ways in which a patent's subject can be proved obvious is by noting that there existed at the time of invention a known problem for which there was an obvious solution encompassed by the patent's claims;

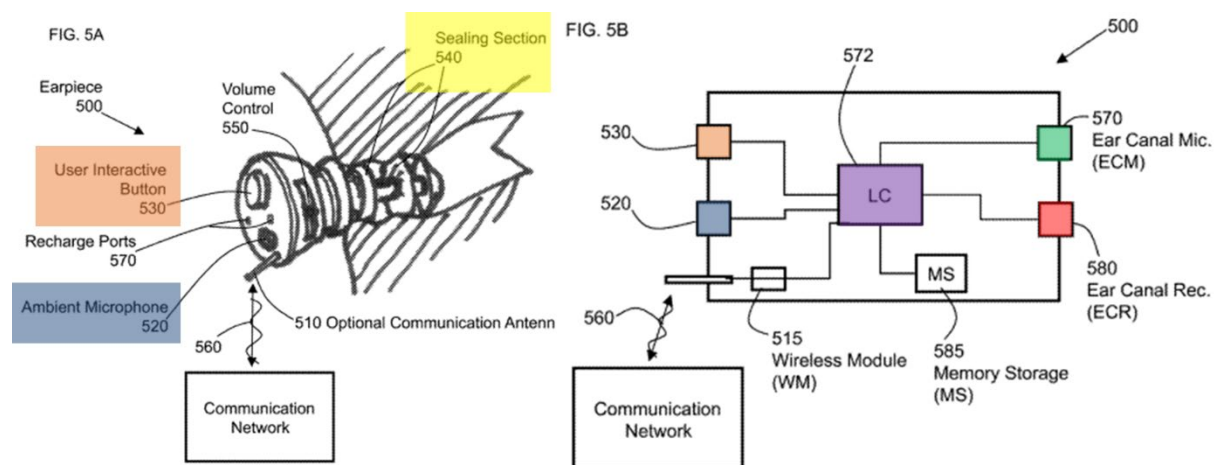
- Any need or problem known in the field of endeavor at the time of invention and addressed by the patent can provide a reason for combining the elements in the manner claimed;
- “Common sense” teaches that familiar items may have obvious uses beyond their primary purposes, and in many cases a person of ordinary skill will be able to fit the teachings of multiple patents together like pieces of a puzzle;
- A person of ordinary skill in the art is also a person of ordinary creativity, and is not an automaton;
- A patent claim can be proved obvious by showing that the claimed combination of elements was “obvious to try,” particularly when there is a design need or market pressure to solve a problem and there are a finite number of identified, predictable solutions such that a person of ordinary skill in the art would have had good reason to pursue the known options within his or her technical grasp; and
- One should be cautious of using hindsight in evaluating whether a claimed invention would have been obvious.

32. I further understand that, in making a determination as to whether the claimed invention would have been obvious to a person of ordinary skill, the Board may consider certain objective factors if they are present, such as: commercial success of products practicing the claimed invention; long-felt but unsolved need; teaching away; unexpected results; copying; and praise by others in the field. These factors are generally referred to as “secondary considerations” or “objective indicia” of nonobviousness. I understand, however, that for such objective evidence to be relevant to the obviousness of a claim, there must be a causal relationship (called a “nexus”) between the claim and the evidence, and that this nexus must be based on a novel element of the claim rather than something in the

prior art. I also understand that even when they are present, secondary considerations may be unable to overcome primary evidence of obviousness (such as motivation to combine with predictable results) that is sufficiently strong.

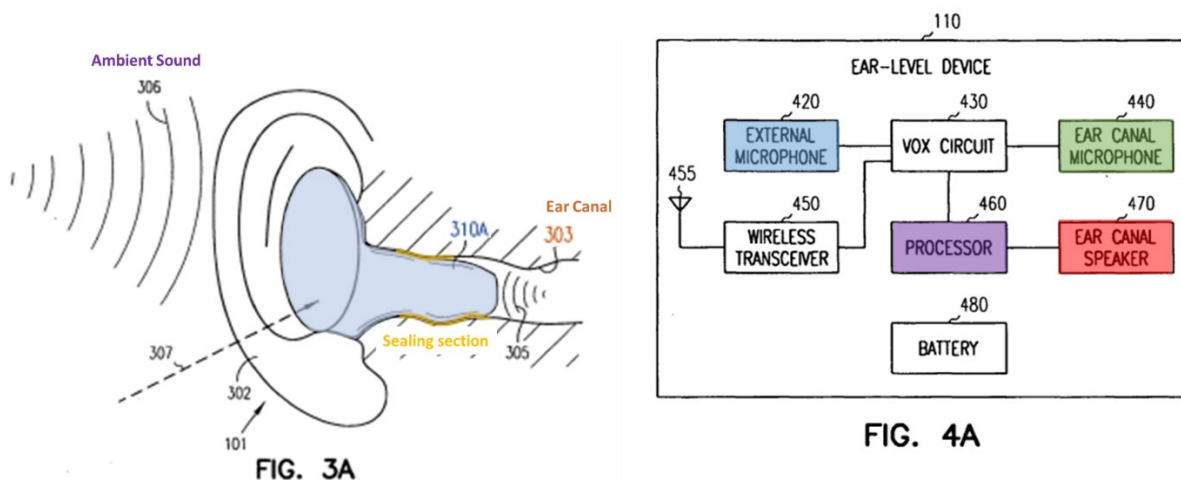
IV. BACKGROUND OF THE TECHNOLOGY

33. As discussed in Section V, the '632 patent is directed to an earpiece device with a **sealing section** that seals an ear canal. '632 patent, '9:15-18. The earpiece device includes a **user interaction element 530**, an **ambient microphone (AM) 520** picking up ambient audio content, an **ear canal microphone (ECM) 540** picking up audio in the ear canal, an **ear canal receiver (ECR) 580** directing audio content to the ear, and a **logic circuit 572** that performs operations related to these components. '632 patent, 9:15-18, 9:32-37. The earpiece performs noise reduction and includes a "voice controlled interface." '632 patent, 11:54-56, 13:43-44.



'632 patent, Figure 5A and 5B.

34. The technology at issue in the '632 patent was developed by others before the '632 patent. For example, Victorian discloses an “ear-level communication system” that takes the form of an “earmold” that “is fully **occluded** to maximize audio efficiency.” Victorian, [0018], [0027]. Victorian’s “ear-level device” includes “**an external microphone 420,**” “**ear canal microphone 440,**” “ear canal speaker 470,” and “processor 460” with “noise reduction circuitry.” Victorian, [0030]. Victorian’s ear-level communication system is illustrated in Figures 3A and 4A below, which I have annotated.



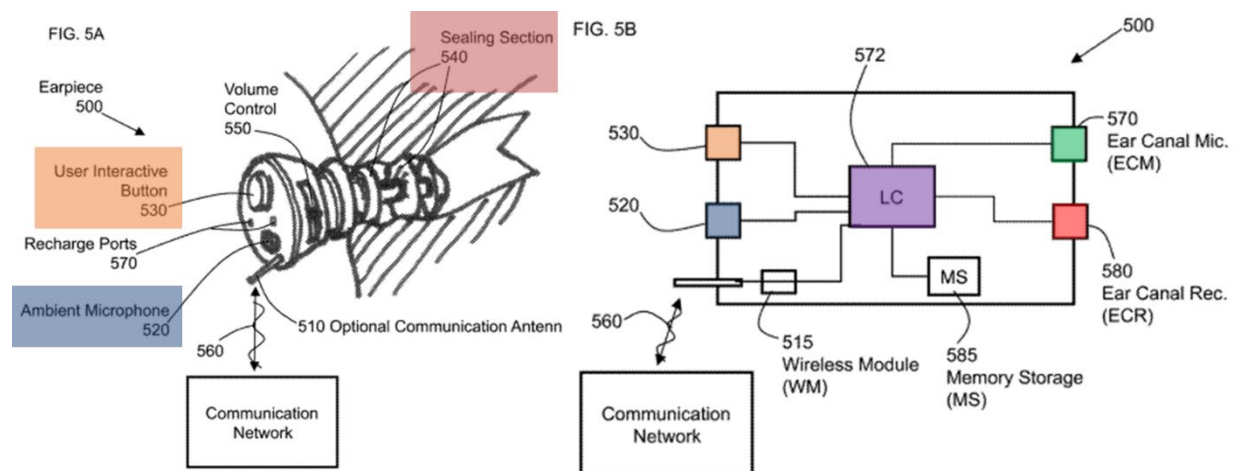
Victorian, Figs. 3A and 4A.

35. In addition, Jannard discloses an “audio device” with “a microphone that receives a voice command.” Jannard, [0316]. Jannard also discloses **button 150**, which acts as an interaction element. Jannard, [0291].

V. THE '632 PATENT

A. '632 Patent Overview

36. The '632 patent describes an earpiece in which “control of audio content” is initiated by a “voice command detected in captured audio.” '632 patent, Abstract. The earpiece is illustrated in Figures 5A and 5B, and includes **sealing section 540**, **ambient microphone 520** “to pick up ambient audio content,” **ear canal microphone 570** to “pick up audio in the ear canal region,” and **ear canal receiver 580** to “direct audio content to the ear drum.” '632 patent, 9:15-44.



'632 patent, Figures 5A and 5B.

37. The earpiece also includes **logic circuit 572**, including “speech synthesis and recognition.” '632 patent, 8:48-52. A voice-controlled interface is activated when an **interaction element, such as a button**, is pressed. '632 patent, 11:49-12:3. After the voice-controlled interface is activated, a user can issue a

voice command to control audio content, such as skipping to the next piece of audio content. '632 patent, 11:57-65. The earpiece also performs noise reduction that can be customized to a user's specific environmental conditions. '632 patent, 13:43-49.

B. The Prosecution History

38. The '632 patent issued from U.S. Pat. App. No. 14/148,749 ("749 Application"), filed January 7, 2014. The '632 patent claims priority to U.S. Provisional Application No. 60/806,769, filed July 8, 2006.

39. During prosecution, the Examiner issued a rejection of claims 1-3 under 35 USC §103 based on Tanaka (US 2003/0130852A1) in view of Rezvani (US 20070165875 A1). According to the Examiner, Tanaka teaches an earpiece, a sealing section, an ambient microphone, a speaker, a memory that stores instructions, a logic circuit that executes the instructions to perform operations, the operations comprising: detecting manual activation, and switching the mode of the earpiece to voice command when manual activation is detected. '632 FH, 174-175. The Examiner stated that Tanaka does not teach a location unit, but that Rezvani does, and it would have been obvious to a person of ordinary skill in the art to modify Tanaka to include at least a GPS module, as taught by Rezvani. *Id.* The Examiner also rejected the dependent claims, including claims 16-18 based on Takana in view of Rezvani and Victorian based on Victorian's disclosure of an ear

canal microphone.

40. In response, Applicant amended Claim 1 to add “an ear canal microphone configured to measure internal audio in the ear canal of the user,” and specify that the sealing section is “configured to be inserted into and seal an ear canal of a user,” the ambient microphone is “configured to measure ambient audio content in the ambient environment near the user’s ear,” the interaction element “is a tactile interface,” the location unit is “configured to determine the location of the earpiece,” and that the logic circuit performs the following operations: “generating active noise reduction using at least one of the measured ambient audio, the internal audio or a combination of both”; “activating a voice controlled interface when” manual activation of the interaction element “is detected”; “detecting a user’s voice from measure ambient audio content; extracting an audio signal of the user’s voice; and analyzing the audio signal to determine a voice command.” ’632FH, 206.

41. Applicant argued that both Tanaka and Rezvani disclose “a headset, not an earpiece,” and that neither Tanaka nor Rezvani teaches “an ambient mic near the user’s ear.” Applicant further argued that Rezvani’s headset is not combinable with an earpiece because “microphone array” “since the placement of microphones is limited in an earpiece (thus difficult for beam forming).” ’632FH, 212.

42. The Examiner then issued a notice of allowance without comment.

VI. LEVEL OF ORDINARY SKILL IN THE ART

43. I have been asked to assume the Claims are entitled to the July 8, 2006 filing date of provisional application No. 60/806,769 for the purposes of this declaration only.

44. I have been informed by Petitioner's counsel that a person having ordinary skill in the art is a hypothetical person who is presumed to have the skill and experience of an ordinary worker in the field at the time of the priority date of the challenged patent.

45. Based on my knowledge and experience in the field and my review of the '044 patent and file history, I believe that a person having ordinary skill in the art in this matter would have had at least a Bachelor's of Science Degree (or equivalent) in electrical engineering, computer engineering, audio engineering or the equivalent, and at least two years of experience with audio devices and/or audio signal processing, such as earphones, hearing aids and/or hearing protectors. A greater amount of education could compensate for fewer years of work experience, and vice versa.

46. Based on my experiences, I have at least the skills of a person having ordinary skill in the art as well as a good understanding of the capabilities of a person having ordinary skill in the art. To be clear, all of my opinions in this

declaration are from the perspective of one of ordinary skill in the art as I have defined it here during the relevant timeframe.

VII. CLAIM CONSTRUCTION

47. I have been instructed that the Challenged Claims of the '632 patent are to be given their ordinary and customary meaning as understood by one of ordinary skill in the art at the time of the alleged invention in light of the specification and the prosecution history pertaining to the patent. I have followed these principles in my analysis set forth in this document.

48. I understand that the parties have not proposed constructions for the '632 patent in the Harman Litigation.

VIII. GROUNDS OF UNPATENTABILITY

49. As I explain above in Section IV, a dual-microphone earpiece with a sealing section in an ear canal that captures and responds to voice commands was well-known in the art.

50. In this section, I will explain why the teachings or combinations of teachings from the following references render obvious each of the Challenged Claims:

Name	Ex.	Publication/ Patent No.	Published/ Issued
Victorian	1007	EP1519625	3/30/2005
Jannard	1008	US2006/0132382	6/22/2006
Karaoguz	1009	US2004/0117846	6/17/2004

Name	Ex.	Publication/ Patent No.	Published/ Issued
Nathan	1010	US2004/0167847	8/26/2004
Camp	1011	US2007/0274531	3/30/2005

I have been instructed to assume that each of these references qualifies as prior art.

51. As I explain below, the Challenged Claims are invalid based upon the following prior art grounds:

- **Ground 1:** Challenged Claims 1-7, 9, and 15-16 of the '632 patent are rendered obvious by Victorian in view of Jannard.
- **Ground 2:** Challenged Claim 8 of the '632 patent is rendered obvious by Victorian in view of Jannard and Karaoguz.
- **Ground 3:** Challenged Claims 10-14 of the '632 patent are rendered obvious by Victorian in view of Jannard and Nathan.
- **Ground 4:** Challenged Claims 16-17 of the '632 patent are rendered obvious by Victorian in view of Jannard and Camp.

52. These Grounds above are not substantively the same as the art or arguments raised in the prosecution of the '632 patent. While the Examiner identified and considered Victorian, the Examiner did not consider it in combination with Jannard, which was not considered during prosecution of the '632 patent. '632FH, 176, 179. And each of these references teaches limitations that the Examiner erroneously believed were missing from the prior art. For example, unlike Tanaka and Rezvani, Jannard teaches a system with voice-activated electronic controls of media of a wearable device (Jannard, [0177], [0315]). Karaoguz teaches a media device that allows a user to scan audio options.

Karaoguz, [0024], [0047]. Nathan teaches an audio system that allows financial stock transactions to be initiated by voice commands. Nathan, [0006]-[0007], [0020]. Camp teaches earbuds that measure and calculate sound pressure levels for a user. Camp, [0018], [0115]. Indeed, the Examiner did not identify any art teaching voice-activated controls that are activated using an interaction element.

IX. SPECIFIC GROUNDS FOR CHALLENGE

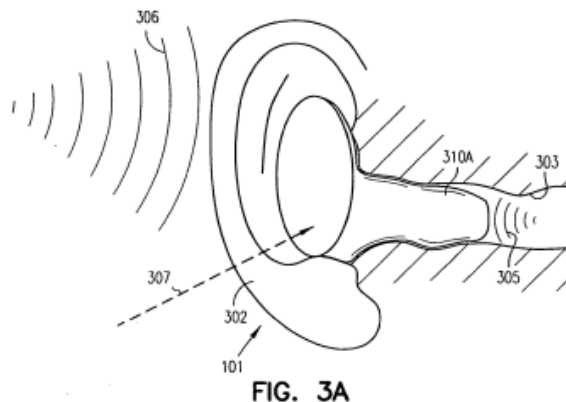
53. As I explain above in Section IV, although the '632 patent purports to have invented a dual-microphone earpiece with noise cancellation and voice command capabilities, such methods were well-known in the art. As described below, the cited prior art provides these teachings.

54. In this section, I explain why claims 1-7, 9, and 15-16 of the '632 patent are rendered obvious by the teachings of Victorian (Ex. 1007) in view of Jannard (Ex. 1008), why claim 8 is rendered obvious by the teachings of Victorian in view of Jannard and Karaoguz (Ex. 1009), why claims 10-14 are further rendered obvious by the teachings of Victorian in view of Jannard and Nathan (Ex. 1010), and why claims 16-17 are rendered obvious by the teachings of Victorian in view of Jannard and Camp (Ex. 1011). I note that, in this discussion, I may focus on the particular language in specific claims, but it should be understood that my analysis applies to the corresponding limitations of all claims containing identical or substantially identical language.

A. GROUND 1: Victorian in view of Jannard (Claims 1-7, 9, 15-16)

1. Overview of Victorian

55. **Victorian** is directed to an “ear-level full duplex audio communication systems for communicating with remote devices. These systems each include one or two ear-level devices each worn in or about an ear, such as in-the-ear (ITE) or behind-the-ear (BTE) devices, wirelessly coupled to a remote device such as a computer, a personal digital assistant (PDA), a cellular phone, a walkie talkie, or a language translator.” Victorian [0016]. The ear-level device may be used as an “ear piece wirelessly extended from the remote device.” Victorian, [0016]. The ear-level device is in the form of an “the earmold [that] is fully occluded to maximize audio efficiency.” Victorian, [0027], Fig. 3A.



Victorian, Figure 3A.

56. An embodiment of Victorian’s ear-level device is illustrated in Figure 4A. Victorian, [0029]-[0030]. The ear-level device includes “an **external microphone 420**” that “picks up ambient sound,” “**ear canal microphone 440**”

that “detects sound from the ear canal,” and “[e]ar canal speaker 470” that plays sound from **external microphone 420** or received from a remote device via “wireless transceiver 450.” Victorian, [0030]. Victorian’s ear-level device also includes **processor 460**, including “noise reduction circuitry.” Victorian, [0030].

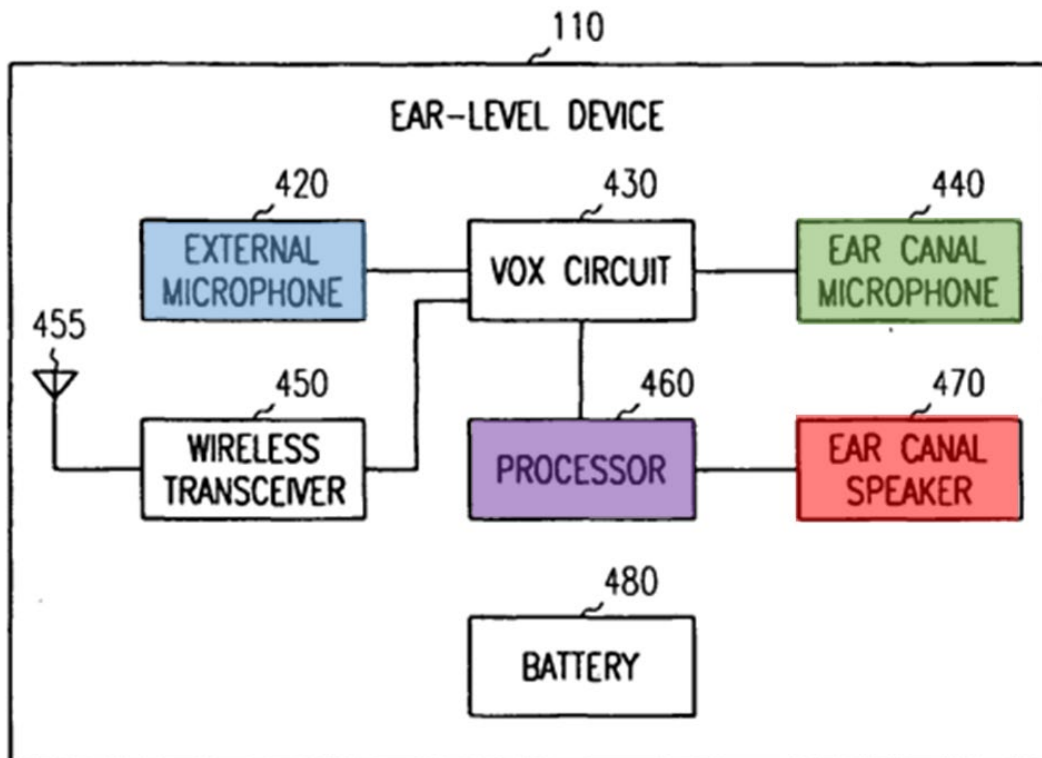


FIG. 4A

Victorian, Figure 4A.

57. **Processor 460** also includes “a speech recognition module to enhance the audio signal received by and/or transmitted from ear-level device 110.”

Victorian, [0030]. Victorian discloses in Fig. 9 how speech recognition is

performed on audio from **ear canal microphone 440** (steps 910-918), the remote device (steps 920-928), and **external microphone 420** (steps 930-938). Victorian, [0050]-[0052]. Victorian explains the details of this process for **external microphone 420**: “Steps 930-938 illustrate the process of detecting a sound from the environment and transmit it to the ear canal. A microphone outside the ear canal detects an ambient sound at 930. If the ambient sound is detected at 931, the VOX of the system is started at 932 if it is not already on. The ambient sound is converted to a second incoming audio signal at 934. The second incoming audio signal is enhanced at 936 with a voice recognition system that improves the signal-to-noise ratio, where the signal is any voice included in the ambient sound, and the noise is any sound other than the voice.” Victorian, [0052].

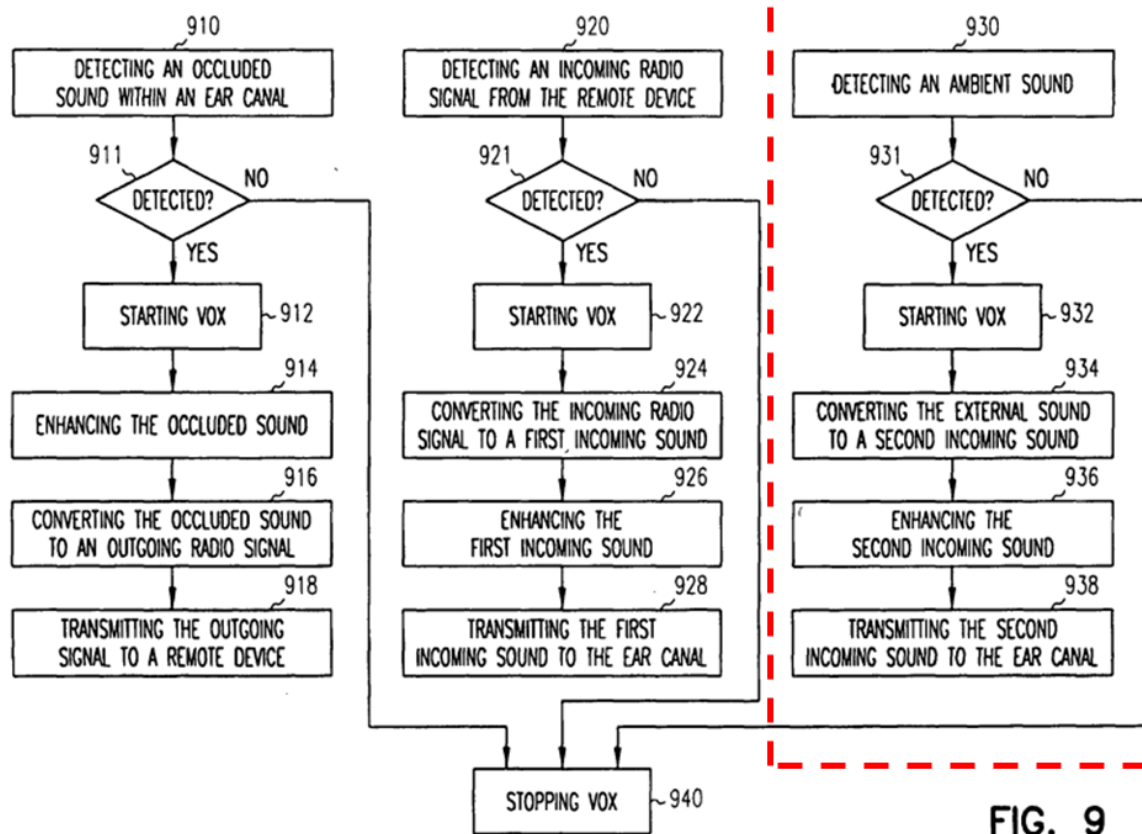


FIG. 9

Victorian, Figure 9.

2. Overview of Jannard

58. Jannard discloses an “[a]udio device 304 generally includes a data port 308, data interface 310, **processor 312**, digital-to-analog converter 314, speaker drivers 316, and speakers 318. In addition, audio device 304 generally also includes a control interface 320, user controls 322, display/indicator drivers 324, display/indicators 326, power module 328, and memory module 330; however, any one or more of these components may be combined.” Jannard, [0304], Fig. 22. The device can include a “wearable. [sic] electronic interface 601 clothing [that]

include[s] headphones, ear phones, a hat, helmet, goggles, mask, visor, headband, hair band, shirt, scarf, sweater, jacket, belt, pants, vest, etc.” Jannard, [0209].

Jannard also discloses “the user controls 322 include a microphone that receives a voice command.” Jannard, [0316]. Jannard also discloses that a voice command mode can be activated by a user depressing a button. “[A] user can initiate an audio menu by depressing the button 150” Jannard, [0291]. Further “one audio menu option can include voice command. For example, when a user chooses the voice command option, the interface electronic 116 can reconfigure to send an AT command to the source device, such as a cellular phone, to accept voice commands directly from the transceiver 114.” Jannard, [0296]. *See also* [0297]. Jannard further discloses the device includes “noise cancellation electronics (not shown) configured to filter wind-generated noise from an audio signal transmitted from the microphone 75.” Jannard, [0165].

59. Jannard’s audio device is connected to a “source device,” such as a “PDA.” Jannard, [0194]. In addition to using voice commands to “control the audio device,” Jannard’s audio device includes “interface electronics 116” that will “recognize speech from the user, convert the speech to AT commands, and control this source device using a standard AT command protocol.” Jannard, [0298], [0315].

60. Jannard teaches using voice control for internet browsing. Jannard

discloses that existing “devices can communicate with a variety of wireless networks for receiving email, phone calls, and/or internet browsing.” Jannard, [0299]. Jannard further discloses that “such a hand-held computer can include a text-to-speech engine. Thus, such a hand-held computer can be used in conjunction with any of the audio devices 10, 10A, 10A’, 10B to allow a user to ‘hear’ emails, or other text documents without the need to hold or look at the device B. Preferably, the hand-held computer includes a further wireless transceiver compatible with at least one of the transceivers 114, 114i. As such, a user can use any of the audio devices 10C, 10D, 10E to ‘hear’ emails, or other text documents without the need to hold or look at the device B.” Jannard, [0299].

61. Moreover, Jannard’s audio device also allows control of audio content (where such control is via user controls 322 including a microphone). “[U]ser controls 322 may be used to fast-forward, skip, cue, play, pause, turn power on or off, rewind, review, adjust volume, balance, tone, bass, or treble, randomize file selection, load a playlist, set a playlist, delete a playlist, repeat playback of all audio files, selected audio files, or a playlist, or perform any other function related to an audio data file.” Jannard, [0317]. Jannard discloses that audio content comes from “[a]udio network 300 [that] includes a content source 302 coupled to an audio device 304 via communications link 306. The content source 302 is any of a variety of information sources, including, but not limited to [...] the Internet, or the

like. The content provided by the content source 302 may be any of a variety of information, including but not limited to, audio files, entertainment, news, media, music, photos, videos, advertising, etc.” Jannard, [0301]. To obtain the audio content, a user provides “a list of preferred information, which includes preferred song names, artist names, or show or program titles. The audio device 304 monitors wireless signals to determine if preferred information is or will be broadcast.” Jannard, [0328]. Audio content is also obtained from “FM or satellite radio.” Jannard, [0327].

3. Motivation to Apply Jannard’s Teachings to Victorian

62. Victorian and Jannard are analogous art. Victorian and Jannard are in the same field as the ’632 patent, which includes personal audio devices. ’632 patent, Abstract; Jannard, [0165], [0265], [0326]; Victorian, [0030], [0040], [0041], [0052]. Victorian and Jannard are reasonably pertinent to alleged problems addressed by the ’632 patent of using sound detection in a personal audio device to improve functionality. *E.g.*, ’632 patent, Abstract (“The interaction element and logic circuit are configured to initiate control of audio content for output from the speaker in response to at least one voice command detected in captured audio.”); Victorian, [0003] (“An ear-level communication system generally includes tethered or wireless headsets with a microphone boom or a microphone within a cord of the headset to detect voice of the person wearing the headsets.”); Jannard,

[0316] (“[T]he user controls 322 include a microphone that receives a voice command.”). For example, both Victorian and Jannard disclose personal wireless audio devices that connect to an audio source device, and use microphones to detect sound to provide, e.g., noise cancellation. Victorian, [0030] (“Ear canal microphone 440 detects sound from the ear canal....Processor 460 converts the sound picked up from the ear canal to an electrical signal to be transmitted to remote device 190, and converts the signals received from remote device 190 and/or external microphone 420 to a sound audible to the person wearing ear-level device 110. In one embodiment, processor 460 includes one or more of amplification circuitry, filtering circuitry, acoustic feedback reduction circuitry, noise reduction circuitry.”); Jannard, [0165] (“[A]udio device 10A” can include noise cancellation electronics (not shown) configured to filter wind-generated noise from an audio signal transmitted from the microphone 75.”) and [0176] (“For example, but without limitation, the source device B can be configured to communicate with other wireless data networks such as via, for example, but without limitation, long-range packet-switched network protocols including PCS, GSM, and GPRS. As such, the audio device 10C can be used as an audio interface for the source device B.”).

63. A person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of using a voice control interface to Victorian’s ear-level

device to improve convenience by advantageously allowing connected devices to be controlled with voice commands. Victorian explains that ear-level devices “allow private, hand-free use” of a communication device such as a cellular phone or PDA but does not specify a particular mechanism for providing such hands-free use. Victorian, [0002], [0016]. It was understood that using voice control to control a device would be one mechanism for providing such “hand-free use” of a connected device, and that this would be more convenient for a user, and make the product more marketable. For example, Ex. 1013 (Stahl) explains that, “[t]o increase the user-friendliness and options for use of devices...and thus to make the devices more attractive, an ever increasing number of devices are so equipped that control of the device is possible by means of acoustic command signals.” Ex. 1013, ¶[0002]. Stahl also states:

As speech recognition systems develop, devices have also become available which may recognize and accept various voice commands as command signals, so that complicated control of such devices is also possible. Such voice-controllable devices are highly convenient, since the operator may operate the respective device without having to use his/her hands. This control method consequently has considerable advantages wherever the operator needs his/her hands for other activities, for instance in the case of control of a car radio, where the operator must not take his/her hands off the steering wheel to change the volume or the channel. In addition, this method is also more generally attractive with regard to device operation, because such voice control enables the man-machine interface (MMI) to be shifted from the hitherto conventional plane of communication with machines, namely operation by buttons and controllers, to the communication plane normal to humans, namely information transfer via speech.

Ex. 1013, ¶[0002]. In addition, Ex. 1017 (Cohen) explains that “[a]ctivities such

as driving occupy the user's hands and eyes. Speech is an ideal solution for accessing services while engaged in such tasks." Ex. 1017, p. 11. This was also recognized to be further useful in smaller devices, which have "very poor user interfaces," such that "Speech is the ideal solution." Ex. 1017, p. 11. Similarly, it was recognized that "[m]any cell phones also include a 'hands-free' or voice command setting that may be utilized when the user is ... unable to manually operate the device." Ex. 1018 (Chu), [0003].

64. In addition, Victorian already includes microphones that detect a user's voice and a processor to run software, and thus, the inclusion of a voice control interface as taught by Jannard would have been straightforward. Victorian, [0040].

65. In this regard, a person of ordinary skill in the art would have been motivated to further apply Jannard's teachings of using voice controls to control the various PDA functionalities taught by Jannard, including "internet browsing" over a connected PDA, and controlling audio content both on the audio device and obtained from other sources. A person of ordinary skill in the art would have recognized that expanding the voice control capability of Victorian's ear-piece device to process voice commands relating to these features would increase the range of devices Victorian's ear-piece device could conveniently control in a hands-free manner and improve the user's experience. For example, by 2006, it

was recognized that “[m]any wireless cellular phones now provide “wireless web” limited browsing capability,” and that “use of these products supports the trend of access to Internet content beyond the PC.” Ex. 1019 (Janik), [0018]. One issue that was recognized with using such smaller devices was “limited display area and awkward methods of user interaction.” *Id.* The complex inputs needed for internet browsing, including multiple characters for a given request, would have taken significant focus to input using a traditional interface on a small device, but would have been easier to provide with a voice interface. Ex. 1019 (Cohen), p. 11. In addition, people often listen to music while performing other tasks, and the ability to control playback of music without needing to look at a device interface would allow a user to perform that other task without being distracted. Ex. 1013 (Stahl), [0002].

66. A person of ordinary skill in the art also would have been motivated to apply Jannard’s teachings of a button activating the voice control interface to Victorian’s ear-level device to prevent unwanted use of voice controls. A person of ordinary skill in the art would have recognized that, in implementing a voice control interface, if the interface is always listening for voice controls, the user may inadvertently cause the device to operate in response to a voice command even though the user was, *e.g.*, just having a conversation with another person. For example, it was well-known that “inadvertent commands” were a problem with

voice interfaces and that voice command devices needed to be programmed to ignore these types of speech. Ex. 1020 (Daum), [0016]. Moreover, a person of ordinary skill in the art would have recognized that using Jannard's button to activate voice controls would reduce unnecessary use of processing resources. Ex. 1015 (Spoken Language Processing), Page 430. For example, it was recognized that "[t]he push-to-talk mode [that] uses a special push event to activate or deactivate speech capturing, which is immune to the potential background noise and can eliminate unnecessary use of processing resources to detect speech events." Ex. 1015 (Spoken Language Processing, Huang, 2001 ("SLP")), Page 430. By using a simple, physical interface element like a button, this would ensure that the user would be aware of when the device is listening for a voice command, and would be able to speak freely at other times. A person of ordinary skill in the art would have recognized that this type of "push-to-talk" feature was common for speech recognition devices at the time. For example, "[i]n order to prevent inadvertent commands being issued to the device and restrict access to its voice control feature," "the voice control feature could be rendered inactive until a specific input signal is received which could be, for example...operation of a tactile sensor incorporated into the device." Daum, [0016]; SLP, Page 430.

67. A person of ordinary skill in the art would have also been motivated to apply Jannard's teachings of a GPS unit to Victorian's ear-level device to

provide directions. For example, Jannard explains “a blind wearer of audio device can employ...a GPS unit for direction-finding.” Jannard, [0267]. In addition, a person of ordinary skill in the art would have understood that, by keeping track of the location of the ear-level device, it would be easier to locate the ear-level device when lost. For example, there was a demand to use GPS to locate and track valuable devices—like an audio device. Ex. 1021 (Fuchs), [0005] (“With the advent of the Global Positioning System (GPS), there is growing demand for mobile devices that can locate and track children, the elderly, emergency situations, tourists, security, valuable assets, and the like.”).

68. A person of ordinary skill in the art also would have been motivated to apply Jannard’s teaching of using an ambient microphone to perform noise reduction to Victorian’s ear piece. Victorian teaches processor 460 includes “noise reduction circuitry,” but does not specify the process used to perform noise reduction. A person of ordinary skill in the art would have looked to a reference like Jannard, which teaches using a signal from the ambient microphone to perform noise reduction. Jannard, [0165]. A person of ordinary skill in the art would have used the ambient audio because that includes the sounds the user wants to reduce with the noise reduction circuitry, and it would not be possible for the device to remove that noise if it does not know what the noise is. For example, it was known that one “important aspect of developing a speech enhancement system is to

accurately assess what information can be assumed about the speech and the background noise. Given a noisy speech signal with no assumptions of the speech or noise, there is little that can be done to enhance the speech signal. ... [T]he more we know about each component, the better we can solve the problem. Depending on the nature of the corrupting noise, some information of the noise may be obtained from the knowledge of the source, or from actual measurements.” Ex. 1024, Lim, Page 21. In this regard, removing noise (i.e., noise reduction) would similarly be performed using measurements of the ambient audio containing the noise so that the device can determine what noise to remove.

69. In addition, a person of ordinary skill in the art would have found the application of this functionality in Victorian’s ear-level device to be straightforward, as Jannard explains that its audio device including all the above features can be implemented in “ear phones.” Jannard, [0209].

70. Thus, a person of ordinary skill in the art would have found it routine, straightforward, and advantageous to apply Victorian’s teachings of an external microphone to Jannard’s teachings of using a microphone that enables voice recognition within ambient noise. The combination would yield the claimed limitations and would predictably work to provide the expected functionality.

4. Invalidity of Claims 1-7, 9, and 15-16 Over Victorian in view of Jannard

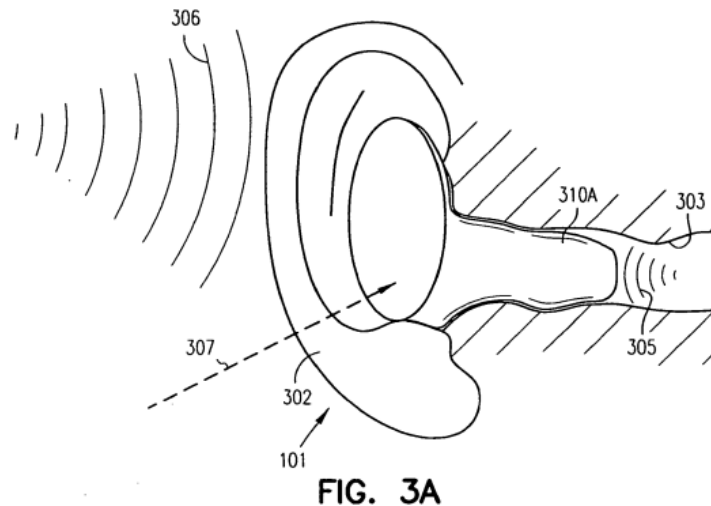
71. As I discuss below, Claims 1-7, 9, and 15-16 are rendered obvious by Victorian in view of Jannard. In addition to my discussion in §IX.A.1-3, the following discussion identifies example disclosures followed by bulleted excerpts that each disclose Claims 1-7, 9, and 15-16.

a. Element [1.pre]

72. **Element [1.pre] recites “An earpiece, comprising.”** This element is met by Victorian.

73. To the extent the preamble is limiting, Victorian discloses an earpiece (e.g., “ear-level device”). *See, e.g.*, Victorian:

- Fig. 3A



- Fig. 4A

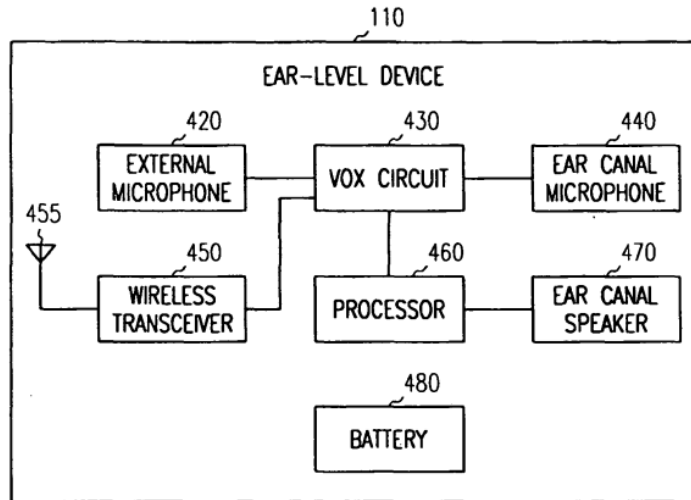


FIG. 4A

- [0008]: “*Ear-level full duplex audio communication systems each include one or two ear attachments devices, such as in-the-ear (ITE) or behind-the-ear (BTE) devices, that wirelessly communicates to a remote device. Such a system allows a hearing impaired individual to communicate with or through the remote device,* such as to talk to another person through a cellular phone. Such a system also functions as an ear piece wirelessly extended from the remote device that allows an individual to privately communicate with or through the remote device without the need of holding the device.”
- [0013]: “FIG. 3A is an illustration of one exemplary exterior configuration of *an in-the-ear (ITE) device* used as the single ear-level device of FIG 1 or one of the pair of ear-level devices of FIG. 2. ... FIG. 4A is a block diagram showing one embodiment of the circuit of ear-level device shown in FIGS. 3A, 3B, or 3C used as the single ear-level device of FIG. 1.”
- [0016]: “This document discusses, among other things, *ear-level full duplex audio communication systems for communicating with remote devices.* These systems each include one or two ear-level devices each worn in or about an ear, such as in-the-ear (ITE) or behind-the-ear (BTE) devices, wirelessly coupled to a remote device such as a computer, a personal digital assistant (PDA), a cellular phone, a walkie talkie, or a language translator. When used as a hearing aid, such a system allows a hearing impaired individual to communicate with or through the remote

device, such as to talk to another person through a cellular phone. When being used as an ear piece wirelessly extended from the remote device, such system allows an individual with normal hearing to privately communicate with or through the remote device without the need of holding the device or wearing any device wired to the remote device.”

- [0045]: “FIG. 7 is a block diagram showing one embodiment of the circuit of *ear-level device* shown in FIG. 6 used as ear-level device 110. The exterior configuration of ear-level device 110 is of the type illustrated as BTE device 610.”
- [0046]: “Ear-level device 110 configured as BTE device 610 differs from ear-level device 110 configured as ITE/ITC/CIC devices 310A/310B/310C primarily in the physical arrangement of device components. By way of example, but not by way of limitation, FIG. 7 illustrates one embodiment in which BTE module 611 includes external microphone 420, processor 460, wireless transceiver 450, antenna 455, VOX circuit 430, and battery 480, and ITE module 612 includes ear canal microphone 440 and ear canal speaker 470. Other possible embodiments depend on size and other considerations that are known to those skilled in the art. In one embodiment, behind-ear portion 611 includes only battery 480 to maximize the duration of use of ear-level device 110 between battery replacements or recharges.”

74. Victorian discloses an “ear-level device” used as an earpiece.

Victorian, [0018]. Victorian further discloses its ear-level device is in the form of an “in-the ear (ITE) device.” Victorian, [0008], [0013], Fig. 3A.

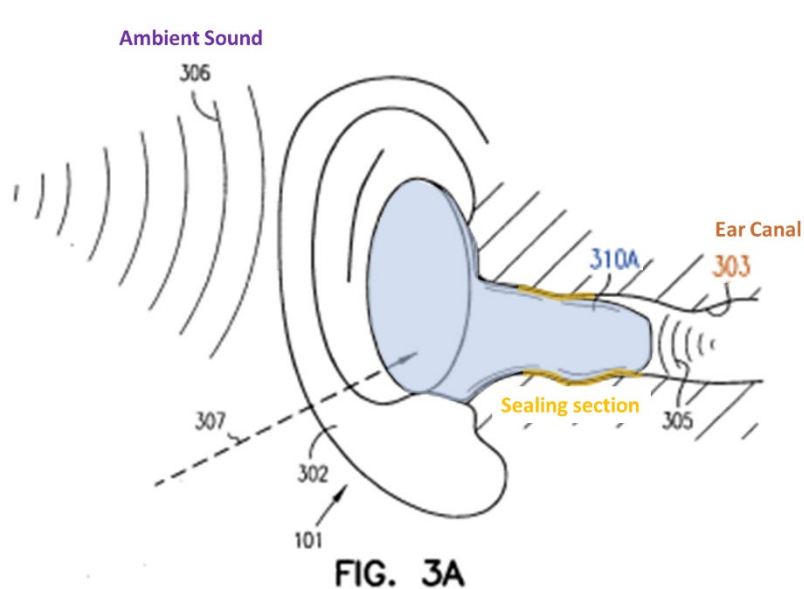
b. Element [1.a]

75. **Element [1.a] recites “a sealing section configured to be inserted into and seal an ear canal of a user.”** This element is met by Victorian.

76. Victorian discloses a sealing section (e.g., portion of “custom earmold” contacting ear, as shown in Fig. 3A) configured to be inserted into and seal an ear canal of a user (e.g., “the earmold is fully occluded to maximize audio

efficiency”). *See, e.g.,* Victorian:

- Fig. 3A (annotated)



- [0004]: “Bi-directional communication has been achieved by a simplex system including a device in the ear canal that both transmits sound to and picks up sound *within the ear canal.*”
- [0027]: “In one embodiment, each of ITE device 310A, ITC device 310B, and CIC device 310C includes a *custom earmold.* In another embodiment, each of ITE device 310A, ITC device 310B, and CIC device 310C includes a *modular earmold.* In one embodiment, *the earmold is fully occluded to maximize audio efficiency.* In another embodiment, the earmold is vented to remove undesirable low frequency sound.”

77. As shown in Figure 3A, Victorian teaches an “earmold” that is “fully occluded” and thus seals the ear canal of the user “to maximize audio efficiency.”

Victorian, [0027].

c. Element [1.b]

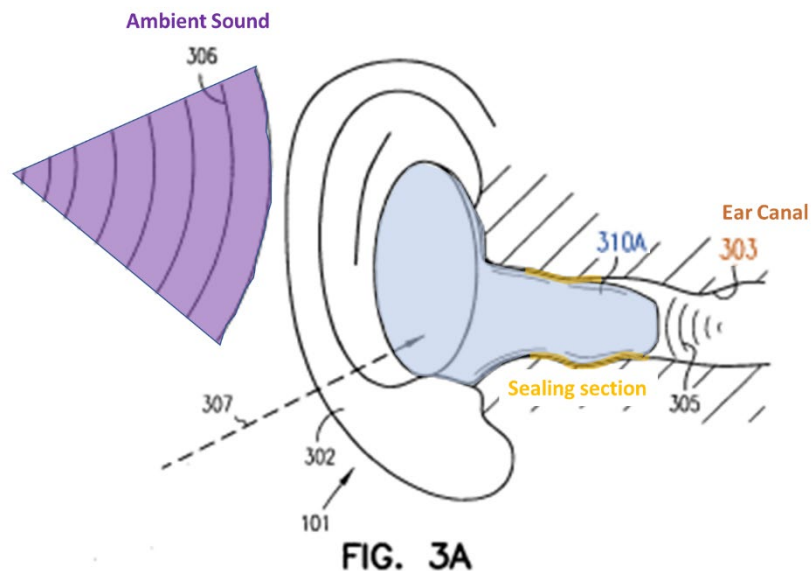
78. **Element [1.b] recites “an ambient microphone configured to**

measure ambient audio content in the ambient environment near the user's ear." This element is met by Victorian.

79. Victorian discloses an ambient microphone (*e.g.*, “external microphone 420”) configured to measure ambient audio content in the ambient environment near the user’s ear (*e.g.*, “[e]xternal microphone 420 picks up ambient sound” and is located at the entrance of the user’s ear as illustrated in Fig. 3A).

See, e.g., Victorian:

- Fig. 3A (annotated)



- Fig. 4A (annotated)

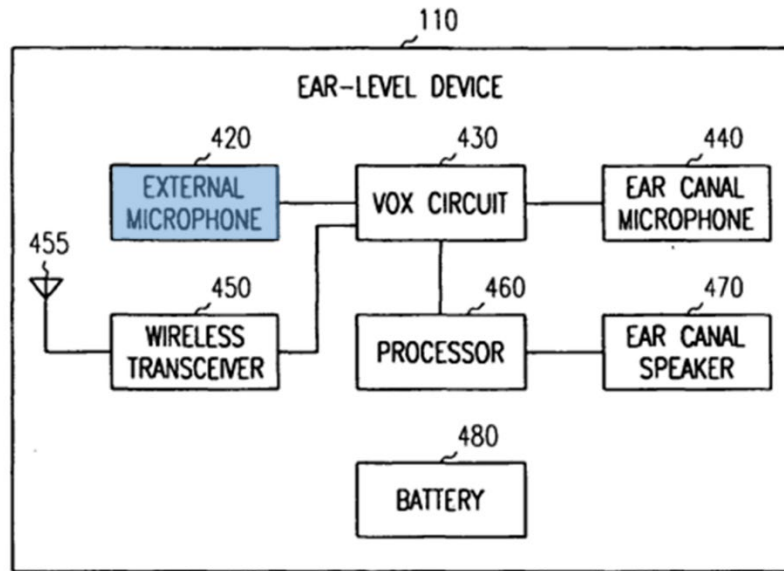


FIG. 4A

- [0030]: “*Ear-level device 110 includes an external microphone 420, an ear canal microphone 440, an ear [sic] canal speaker 470, a processor 460, a wireless transceiver 450, an antenna 455, a voice operated exchange (VOX) circuit 430, and a battery 480. External microphone 420 picks up ambient sound such that a hearing impaired person wearing ear-level device 110 is not “isolated” when communicating to remote device 190. . . . VOX circuit 430 activates a major portion of ear-level device 110 only when sound is detected by at least one of external microphone 420, ear canal microphone 440, and wireless transceiver 450. In one embodiment, VOX circuit 430 includes a voice-controlled switch connecting between battery 480 and the portions of ear-level device 110 that are voice activated. The switch is driven by the sound detected by at least one of external microphone 420, ear canal microphone 440, and wireless transceiver 450. It is turned on when sound detected by external microphone 420 exceeds a predetermined threshold, when sound detected by ear canal microphone 440 exceeds another predetermined threshold, and/or when wireless transceiver 450 receives the Incoming wireless signal representing the remote sound.*”
- [0038]: “When the person wearing ear-level device 110 speaks, both *external microphone 420 and ear canal microphone 440 detect the*

voice. The voice as detected by ear-level device 110 is first transmitted to remote device 190.”

80. See my discussion for element [1.a] above (§IX.A.4.b).

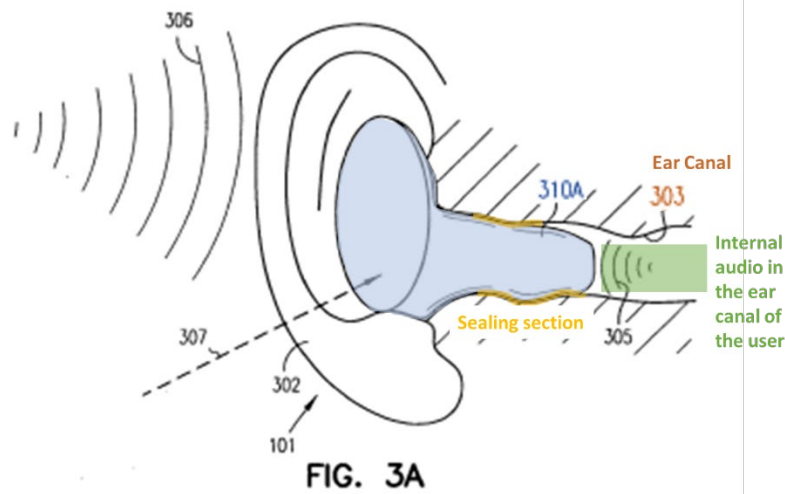
81. Victorian further teaches “external microphone 420 picks up ambient sound,” and that external microphone 420 is located at the entrance of the ear canal. Victorian, [0030], Fig. 3A. Therefore, the ambient sound picked up by external microphone 420 is near the user’s ear.

d. Element [1.c]

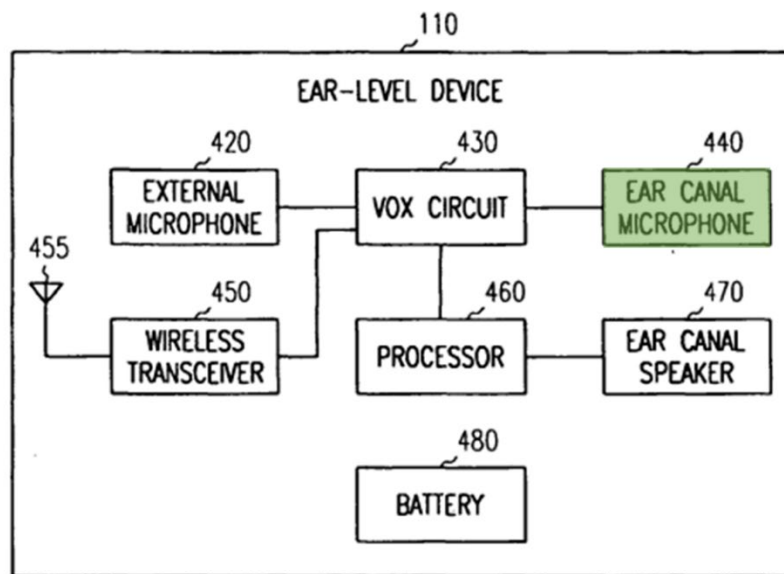
82. **Element [1.c] recites “an ear canal microphone configured to measure internal audio in the ear canal of the user.”** This element is met by Victorian.

83. Victorian discloses an ear canal microphone (*e.g.*, “microphone for use in an ear canal”) configured to measure internal audio in the ear canal of the user; (*e.g.*, “receives an occluded sound from within the ear canal”). *See, e.g.*, Victorian:

- Fig. 3A (annotated)



- Fig. 4A (annotated)



- [0004]: “Bi-directional communication has been achieved by a simplex system including *a device in the ear canal* that both transmits sound to and *picks up sound within the ear canal*. The device is tethered to a push-to-talk (PIT) system such as a walkie talkie. It is not full duplex and therefore requires interaction by the user in order to switch from listen mode to talk mode.”

- [0010]: “In a preferred embodiment, a system includes an ITE module and a BTE module attached to the ITE module. ***The ITE module includes a microphone for use in an ear canal. The microphone receives an occluded sound from within the ear canal.***”
- [0011]: “In a preferred embodiment, a system includes a first ear-level device and a second ear-level device for use in two opposite ears. The first ear-level device includes a microphone, a first processor, and a wireless transmitter. ***The microphone is for use in one ear canal to receive an occluded sound from within that ear canal.***”
- [0030]: “Ear-level device 110 includes an external microphone 420, ***an ear canal microphone 440***, an ear canal speaker 470.... ***Ear canal microphone 440 detects sound from the ear canal. The sound includes primarily speech of the person wearing ear-level device 110.***...In one embodiment, ***ear canal microphone 440*** and ear canal speaker 470 are implemented as one physical device. Processor 460 converts the ***sound picked up from the ear canal*** to an electrical signal to be transmitted to remote device 190, and converts the signals received from remote device 190 and/or external microphone 420 to a sound audible to the person wearing ear-level device 110. In one embodiment, as discussed with respect to FIG. 4B below, VOX circuit 430 gates or attenuates one or more ***sounds detected by*** external microphone 420, ***ear canal microphone 440***, and wireless transceiver 450, to eliminate or reduce echo and ringing caused by the loop formed by ear canal speaker 470 and ear canal microphone 440.”

84. See my discussion for element [1.a] above (§IX.A.4.b).

85. Victorian discloses an “ear canal microphone 440” that “detects sound from the ear canal,” and thus measures internal audio in the ear canal of the user. Victorian, [0030], Fig. 3A. A person of ordinary skill in the art would have understood that to “detect” sound, a microphone would have to measure the sound. For example, microphones operate by detecting air pressure changes caused by noise. Microphones typically have “a diaphragm, whose movement in response to

sound pressure varies the parameters of an electrical system (a variable-resistance conductor, a condenser, etc.), producing a variable voltage that constitutes the microphone output.” SLP, Page 491. A microphone detects sound by measuring pressure changes caused by noise. In this regard, Victorian further discloses the detected sound is compared to “a predetermined threshold,” confirming ear canal microphone 440 measures the sound. Victorian, [0030].

e. Element [1.d]

86. **Element [1.d] recites “a speaker.”** This element is met by Victorian.

87. Victorian discloses a speaker (*e.g.*, “ear canal speaker”). *See, e.g.*,

Victorian:

- Fig. 4A (annotated)

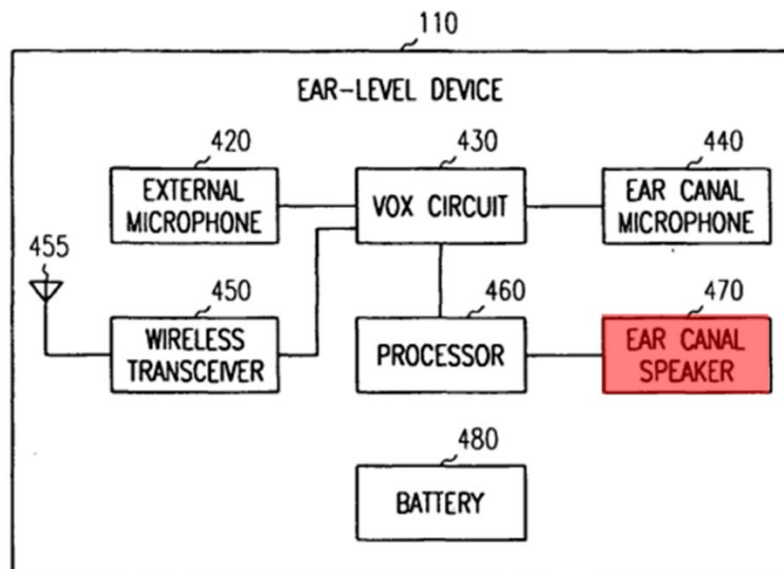


FIG. 4A

- [0030] “*Ear canal speaker 470 transmits sound received from remote device 190 and/or external microphone 420 to the same ear canal.* In one embodiment, ear canal microphone 440 and ear canal speaker 470 are implemented as one physical device.”
- [0031] “Because ear canal microphone 440 is capable of detecting *the sound delivered to the ear canal from ear canal speaker 470*, there is a need to avoid echo and/or ringing. For example, a remote sound from remote device 190 delivered to the ear canal may be detected as an occluded sound and transmitted back to remote device 190, causing echo. The gating modules function in a coordinated manner to eliminate or reduce such echo and/or ringing caused by breaking, or applying resistance into, the audio loop formed by ear canal speaker 470 and ear canal microphone 440.”
- [0046] “By way of example, but not by way of limitation, FIG. 7 illustrates one embodiment in which BTE module 611 includes external microphone 420, processor 460, wireless transceiver 450, antenna 455, VOX circuit 430, and battery 480, and ITE module 612 includes ear canal microphone 440 and *ear canal speaker 470.*”

f. Element [1.e]

88. **Element [1.e] recites “an interaction element, wherein the interaction element is a tactile interface.”** This element is met by Victorian in view of Jannard.

89. Jannard discloses an interaction element (e.g., “button”), wherein the interaction element is a tactile interface (e.g., “user can depress the button 150”).

See, e.g., Jannard:

- [291]: “For example, a user can initiate an audio menu by *depressing the button 150* (FIGS. 10-12) ...Thus, *when the user pushes the button 150* the first time, the user will hear the words “initiate a call,” ...”

- [292]: “[T]he user can depress either the forward or rearward portions 156, 158 of the button 150 so as to “scroll” through other audio menu options.”
- [293]: “[T]he user can depress the button 150 to initiate the audio menu, then “scroll” to the phonebook by depressing the portions 156 or 158 until the user hears the word “phonebook” in the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E. *Once the user hears the word “phonebook,” the user can depress the button 150 again to enter the phonebook.* Thereafter, the user can depress the portions 156, 158 to “scroll” through phonebook entries.”
- [0296]: “*Once the user hears the word “reply” the user can depress the button 150 to enter a reply mode.*”
- [0257]: “[T]he button 152 can be depressed inwardly on a forward end 156 of the button 152, thereby causing the button 152 to pivot or “rock” about the pivot axis 154. Additionally, the button 152 can be pressed at a rearward end 158, thereby causing the button 152 to pivot about the pivot axis 154 in the opposite direction.”
- [0316]: “In one embodiment, the user controls 322 include a **button, dial, switch, lever, sensor, touchpad,** microphone, and/or any other input device that may be used by a user to control the audio device 304.”

90. Jannard teaches the “user can depress the button 150,” e.g., to activate a “voice command option.” Jannard, [0292], [0296]. A person of ordinary skill in the art would have understood a button 150 that can be depressed is tactile. *See, e.g.,* Daum, [0014] (“[T]he tactile sensor actuating the switch could be, for example, a button.”).

91. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a button to activate a voice control interface to Victorian’s in-ear device.

g. Element [1.f]

92. **Element [1.f] recites “a location unit configured to determine the location of the earpiece.”** This element is met by Victorian in view of Jannard.

93. Victorian discloses the earpiece (*see* [1.pre]).

94. Jannard discloses a location unit (*e.g.*, “GPS unit”) configured to determine the location of an earpiece (*e.g.*, “determines position or location” of the audio device). *See, e.g.*, Jannard:

- [0267]: “In another embodiment, a further modification of the audio devices 10, 10A, 10B, 10C, and 10D is provided wherein ***the audio devices can include*** and/or communicate with ***a variety of sensors....*** A blind wearer of ***audio device*** can employ a distance sensor to determine distance to surrounding objects, for example, or ***a GPS unit for direction-finding.***”

95. Jannard further teaches its audio device includes “a variety of sensors,” including a “GPS unit for direction-finding.” Jannard, [0267]. A person of ordinary skill in the art would have understood that a GPS unit in the audio device determines the location of the audio device. *See* Jannard, [0198] (disclosing the use of “GPS” to “determine[] position or location”); *see also* Fuchs, [0039] (disclosing a “method and apparatus for locating a mobile device using “global positioning system (GPS) processing elements”).

96. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a GPS unit to Victorian’s in-ear device.

h. Elements [1.g] and [1.h]

97. Elements [1.g] and [1.h] recite “a memory that stores instructions; and a logic circuit configured to execute the instructions to perform operations, the operations comprising.” These elements are met by Victorian in view of Jannard.

98. Victorian discloses a memory that stores instructions (e.g., memory associated with and storing instructions for “processor 460”) and a logic circuit (e.g., processor 460) configured to execute the instructions to perform operations (e.g., executing instructions to perform operations (see [1.h.1]-[3], [4.g]-[9], [16])).
See, e.g., Victorian:

- Fig. 4A (annotated)

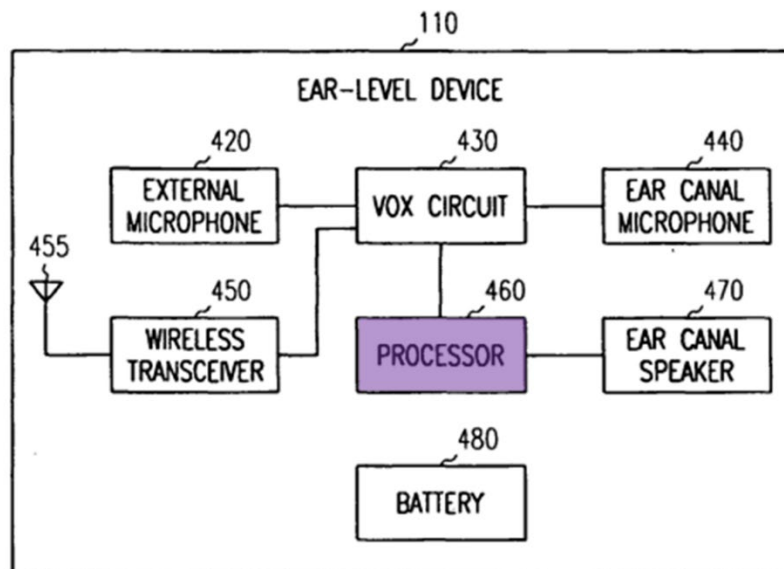


FIG. 4A

- [0030]: “Ear-level device 110 includes an external microphone 420, an ear canal microphone 440, an ear canal speaker 470, **a processor 460**, a wireless transceiver 450, an antenna 455, a voice operated exchange (VOX) circuit 430, and a battery 480. [...] ***Processor 460 converts the sound picked up from the ear canal to an electrical signal to be transmitted to remote device 190, and converts the signals received from remote device 190 and/or external microphone 420 to a sound audible to the person wearing ear-level device 110. In one embodiment, processor 460 includes one or more of amplification circuitry, filtering circuitry, acoustic feedback reduction circuitry, noise reduction circuitry, and tone control circuitry, among other circuits performing signal processing functions as known in the art. In one embodiment, processor 460 includes a speech recognition module to enhance the audio signal received by and/or transmitted from ear-level device 110. Wireless transceiver 450 and antenna 455 form a telemetry interface supporting telemetry link 195 between ear-level device 110 and remote device 190. Wireless transceiver 450 includes a wireless transmitter and a wireless receiver. The wireless transmitter receives the electrical signal representing the sound picked up from the ear canal from processor 460 and transmits an outgoing wireless signal representing the same sound to remote device 190. The wireless receiver receives an incoming wireless signal representing a remote sound from remote device 190 and transmits an electrical signal representing the remote sound to processor 460.***”

99. See my discussion for elements [1.h.1]-[3], [4.g]-[9], and [16] below (§IX.A.4.i-o, p-x, z).

100. Victorian discloses “processor 460” performing operations including “noise reduction,” and processing “sound picked up from the ear canal,” “signals received from remote device 190,” and/or signals received from “external microphone 420.” Victorian, [0030].

101. A person of ordinary skill in the art would have understood that processor 460 has associated memory that stores instructions, and processor 460

executes those instructions to perform operations. At the very least, a person of ordinary skill in the art would have found it obvious to include instructions for processor 460 in memory coupled to processor 460. A processor runs using instructions, and those instructions are stored in memory for the processor to access those instructions. A person of ordinary skill in the art would have recognized that alternatives to storing the instructions in memory in the audio device (e.g., on a separate device) would have made the device more complex and slower as the instructions would need to be sent to the processor from the separate device. The transmission of instructions from the separate device to the ear-level device are not instantaneous and would add to the processing time for the audio device to execute those instructions.

102. To the extent further disclosure is required, Jannard discloses a memory that stores instructions (e.g., memory storing “software”) and a logic circuit (e.g., “processor 312”) configured to execute the instructions to perform operations (e.g., “controls the overall function and operation of audio device 304” using “software executable by the processor 312”). *See, e.g., Jannard:*

- Fig. 22 (annotated)

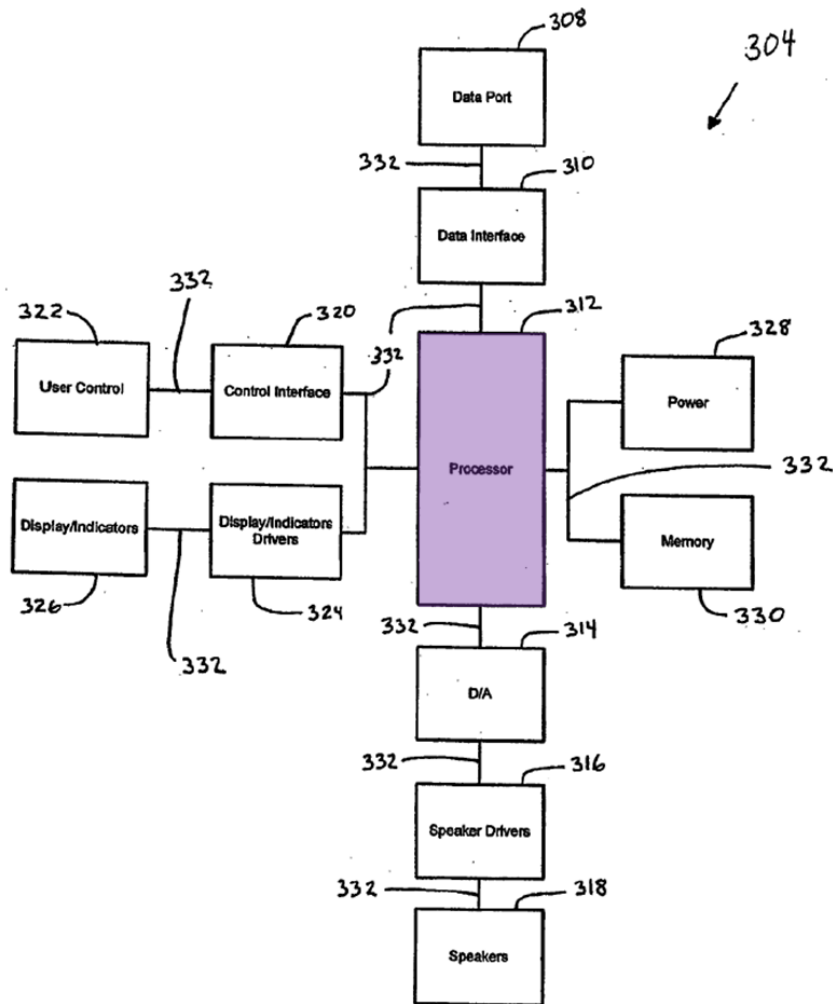


Figure 22

- [0135]: “The device 286 can be configured to store and playback any desired type of electronic audio and/or video file. In the illustrated embodiment, the *device 286 includes a memory*, an amplifier, and a processor. *The memory, amplifier, and the processor are configured to operate together to function as an audio storage and playback system.* For example, the audio storage and playback system can be configured to store MP3 files in a memory and *to playback the MP3 files through the speakers 14A, 16A’.*”
- [0289]: “Further, the interface 116 can include a processor and *a memory for providing added functionality.* For example, the interface 116 can be configured to allow a user to control the cooperating wireless device, such as a cell phone.”

- [0307]: “Data interface 310 couples data port 308 with processor 312. In one embodiment, *data interface 310 is a memory buffer for storing information or content received via data port 308 until it is processed by processor 312. Processor 312 controls the overall function and operation of audio device 304*, and couples directly or indirectly to the various electronic components of the audio device 304, as described herein. In one embodiment, *processor 312 is a digital signal processor (DSP), firmware, microprocessor*, microcontroller, field-programmable gate array (FPGA), and/or an application-specific integrated circuit (ASIC). Processor 312 may also be upgradeable. For example, in one embodiment, *processor 312 is firmware, and software executable by the processor 312 may be changed, uploaded, downloaded, deleted, and/or modified.*”
- [0315]: “The *control interface 320 generally includes a buffer*, register, pre-processor, transistor, resistor and/or other electronic circuit *to enable the processor 312 to receive commands from a user via the user controls 322*. In one embodiment, the *control interface 320 is integrated with the processor 312.*”

103. Jannard teaches that “software executable by the processor 312 may be changed, uploaded, downloaded, deleted, and/or modified.” Jannard, [0307]. Software is implemented as instructions to perform operations. Thus, a person of ordinary skill in the art would have understood that such software is implemented with computer instructions stored in memory. This is further confirmed, for example, by the *Computer Dictionary*, which defines “software” to be “[c]omputer programs; instructions that cause the hardware—the machines—to do work.” EX. 1022 at 365. Jannard further teaches processor 312 “controls the overall function and operation of audio device 304.” Jannard, [0307].

104. A person of ordinary skill in the art would have been motivated to

apply Jannard’s express teachings of a processor executing software stored in memory to implement Victorian’s in-ear device and processor.

i. Element [1.h.1]

105. **Element [1.h.1] recites “generating active noise reduction using at least one of the measured ambient audio, the internal audio or a combination of both.”** This element is met by Victorian in view of Jannard.

106. Victorian discloses generating active noise reduction (*e.g.*, “noise reduction circuitry”) and the measured ambient audio (*e.g.*, “ambient sound” (*see* [1.b])). *See, e.g.*, Victorian:

- [0030]: “In one embodiment, ***processor 460 includes*** one or more of amplification circuitry, filtering circuitry, acoustic feedback reduction circuitry, ***noise reduction circuitry***, and tone control circuitry, among other circuits performing signal processing functions as known in the art.”

107. See my discussion for element [1.b] above (§IX.A.4.c).

108. Victorian further teaches processor 460 includes “noise reduction circuitry.” A person of ordinary skill in the art would have understood the noise reduction is performed based on instructions executed by processor 460. A processor operates by running instructions, and thus processor 460 would have included instructions that cause the noise reduction circuitry to perform noise reduction. For example, “software includes instructions for execution on a digital processor to perform all the steps.” Ex. 1023 (“Gauger”), [0080].

109. Jannard discloses generating active noise reduction (e.g., “noise cancellation function”) using measured ambient audio (e.g., “noise cancellation electronics” “to filter wind-generated noise from an audio signal transmitted from the microphone 75”).

- [0165]: “Further, the audio device 10” can include noise cancellation electronics (not shown) configured to filter wind-generated noise from an audio signal transmitted from the microphone 75.”
- [0326]: “In one embodiment, the audio device 304 includes dual automatic equalization. The audio device 304 may also include static and/or noise correction, and/or *active* or passive *noise cancellation*.”

110. Jannard teaches “noise cancellation electronics” that “filter wind-generated noise from an audio signal transmitted from the microphone 75.”

Jannard, [0165]. See [1.b].

111. A person of ordinary skill in the art would have understood that “wind-generated noise” is “ambient noise.” Wind-generated noise occurs outside of the user’s ear canal and, thus, is measured as “ambient audio.” Moreover, a person of ordinary skill in the art would have understood, and at least found it obvious, that the removal of wind-generated noise uses the signal from microphone 75. For example, in order to remove the wind-generated noise, the noise reduction processing will recognize what the wind-generated noise to be removed is. This knowledge of the wind-generated noise comes from the microphone that detects the noise.

112. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of performing active noise reduction using measured ambient audio to Victorian’s ear-level device. Moreover, a person of ordinary skill in the art would have understood that, when these teachings are applied to Victorian, they would be implemented using Victorian’s noise reduction circuitry in processor 460.

j. Element [1.h.2]

113. **Element [1.h.2] recites “activating a voice controlled interface when a manual activation of the interaction element is detected.”** This element is met by Victorian in view of Jannard.

114. Jannard discloses activating a voice controlled interface (*e.g.*, “interface electronics 116” with “voice command option”) when a manual activation of the interaction element is detected (*e.g.*, when a user “depress[es] the button 150”). *See, e.g.*, Jannard:

- [0291]: “For example, ***a user can initiate an audio menu by depressing the button 150*** (FIGS. 10-12). Upon initiation of the audio menus, the interface electronics 116 can send an audio signal to the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E including a humanoid voice.”
- [0292]: “[T]he user can depress either the forward or rearward portions 156, 158 of the button 150 so as to “scroll” through other audio menu options. For example, other audio menu options can include, for example, but without limitation, phonebook, email, clock, voice commands, and other menu options typically available on cellular phones and/or personal audio devices such as MP3 players.”

- [293]: “*As an illustrative, but non-limiting example*, if a user wishes to access the phonebook, *the user can depress the button 150 to initiate the audio menu*, then “scroll” to the phonebook by depressing the portions 156 or 158 until the user hears the word “phonebook” in the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E. *Once the user hears the word “phonebook,” the user can depress the button 150 again to enter the phonebook.*”
- [0296]: “[A]fter the text-to-speech engine finishes “reading” the email to the user, the interface device 116 can enter another audio menu. Such an audio menu can include a reply option, a forward option, or other options. If a user wishes to reply, the user can “scroll” until the user hears the word “reply.” Once the user hears the word “reply” the user can depress the button 150 to enter a reply mode. As noted above, these types of commands can be issued using an AT command protocol, to which the source device will also be configured to respond. As noted above, one audio menu option can include voice command. For example, when a user chooses the voice command option, the interface electronic 116 can reconfigure to send an AT command to the source device, such as a cellular phone, to accept voice commands directly from the transceiver 114. Thus, as the user speaks, the audio signal is directed to the source device, which in turn is configured to issue audio indicators back to the user, through the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E, to guide the user through such a voice command.”
- [0297]: “For example, if a user chooses a *voice command option*, the user could issue commands such as, for example, but without limitation, “phonebook” or “call alpha.” With a source device such as a cellular phone, that has a speech recognition engine and that is properly trained to recognize the voice of the user, the user can automatically enter the phonebook mode or directly call the phonebook listing “alpha,” of course, as is apparent to one of ordinary skill in the art, such a voice command protocol could be used to issue other commands as well.”
- [0298]: “In another alternative, the interface electronics 116 can include a speech recognition engine and audio menus. In this alternative, the interface electronics 116 can recognize speech from the user, convert the speech to AT commands, and control this source device using a standard AT command protocol.”

115. See my discussion for element [1.e]. above (§IX.A.4.f).

116. Jannard further teaches scrolling through audio menu options until a “voice command option” is identified, at which point “the user can depress the button 150” to initiate the use of voice commands. Jannard, [0296]. A “speech recognition engine” of “interface electronics 116” processes voice commands. Jannard, [0187], [0298]. Moreover, the “interface electronics 116” are connected to a microphone. Jannard, [0187]. This allows the device to operate as an audio interface to provide voice commands that are processed and sent to a device such as a “cellular phone.” Jannard, [0296].

117. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface activated by a button to Victorian’s ear-level device. A person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice control interface to Victorian, the interface would be implemented in processor 340. For example, Jannard discloses interface electronics 116 “can include a processor and a memory for providing added functionality,” including to support a “hands-free protocol,” which a person of ordinary skill in the art would have understood refers to Jannard’s voice control processing. Jannard, [0289]. Moreover, Victorian teaches “processor 460 includes a speech recognition module,” and a person of ordinary skill in the art would have found it straightforward to include all other functionality for voice control processing together in processor 460. Speech

recognition involves pattern recognition based on statistical models. The signal is broken into windows, acoustic features are extracted from the windows, then those are compared statistically against learned models and sequenced (using methods like Hidden Markov Models) to determine what sequence of models is best fit to the sequence of acoustic features. SLP, Page 10 (“Chapter 8 introduces hidden Markov models, the most prominent technique used in modern speech recognition systems.”), 375 (“The hidden Markov model (HMM) is a very powerful statistical method of characterizing the observed data samples of a discrete-time series.”). Because of the variability involved in the statistical analysis used in this process, it is generally performed using software instructions instead of hard coded in hardware. For example, when computing the forward and backward probabilities in HMM calculations with a sufficiently large sequence length, “the dynamic range of these probabilities will exceed the precision range of essentially any machine. Thus, in practice, it will result in underflow on the computer if probabilities are represented directly.” SLP, Page 402. Software advantageously makes these calculations practical where a hard-coded hardware implementation would be insufficient.

118. The voice control feature takes recognized words (which would be constrained to be allowable terms for the menu system) and uses that to activate the software to carry out the corresponding commands. That control needs to be

done in software also, and it is simpler and more efficient to do both the speech recognition and the activation of software to carry out corresponding commands in a single processor.

k. Element [1.h.3]

119. **Element [1.h.3] recites “detecting a user’s voice from measured ambient audio content.”** This element is met by Victorian in view of Jannard.

120. Victorian discloses detecting a user’s voice from a measured ambient audio content (*e.g.*, detection that occurs in “voice recognition system” used on “ambient sound” including “the voice” of “the person wearing ear-level device 110”). *See, e.g.*, Victorian:

- [0030]: “In one embodiment, ***processor 460 includes a speech recognition module*** to enhance the audio signal received by and/or transmitted from ear-level device 110.”
- [0038]: “When the person wearing ear-level device 110 speaks, both external microphone 420 and ear canal microphone 440 detect the voice.”
- [0052]: “Steps 930-938 illustrate the process of detecting a sound from the environment and transmit it to the ear canal. A microphone outside the ear canal detects an ambient sound at 930. ***If the ambient sound is detected at 931, the VOX of the system is started at 932 if it is not already on. The ambient sound is converted to a second incoming audio signal at 934. The second incoming audio signal is enhanced at 936 with a voice recognition system that improves the signal-to-noise ratio,*** where the signal is any voice included in the ambient sound, and the noise is any sound other than the voice.”

121. Victorian teaches “a voice recognition system” that receives “ambient sound” that has been converted into a “second incoming audio signal.” Victorian,

[0054]. The “ambient sound” is captured by external microphone 420 and includes the voice of “the person wearing ear-level device 110.” Victorian, [0040]. A person of ordinary skill in the art would have understood the person wearing the device is a user. A person of ordinary skill in the art would also have understood that, as part of the “voice recognition,” the user’s voice is detected. To perform any voice recognition processing, an incoming audio signal is detected. For example, speech recognition models use different coding for silence compared to speech such that speech is detected when the model shows the received audio is not silence, and instead corresponds to speech. SLP, Page 398 (“If you have the number of states depending on the duration of the signal, you may want to use 15 to 25 states for each second of speech signal. One exception is that, for silence, you may want to have a simpler topology. This is because silence is stationary, and one or two states will be sufficient.”).

122. To the extent further disclosure is required, Jannard discloses detecting a user’s voice from measured ambient audio content (*e.g.*, detection of voice that occurs upon “recogniz[ing] speech from the user”). *See, e.g.*, Jannard:

- [0265]: “In another embodiment, a further modification of the audio devices 10, 10A, 10B, 10C, and 10D is provided wherein the audio devices include at least two banks of microphones, with one bank acting as a speaker of received and ***one bank providing an ambient noise-cancellation function....In a further embodiment, the microphone banks can be arranged in an array to be used in conjunction with algorithms to discern, reduce, and/or eliminate noise for the purpose of voice recognition. For example, in one embodiment, such microphone***

banks can include ASIC-based noise-canceling technology, such as is available in chips from Andrea Electronics Corporation (AEC), to enable voice recognition in ambient noise up to about 130 Db or more.

In another embodiment, microphone banks can be arranged in any suitable combination of linear or non-linear arrays to be used in conjunction with algorithms to discern, reduce, and/or eliminate noise in any suitable manner. *In another embodiment, audio/proximity sensors can advantageously trigger the appropriate functionality in a specific bank.* In another embodiment, a noise-canceling microphone can be provided in connection with a cord or other microphones described above. For example, without limitation, a series of miniature microphones can be supported down a cord from the audio device, separated by desired distances, and aimed in different directions. In another implementation, one or more of the microphones can be for verbal input from the user, and one or more others of the microphones, or the same microphone, can also be for noise-cancellation purposes.”

- [0298]: “In another alternative, the interface electronics 116 can include a *speech recognition engine* and audio menus. In this alternative, the interface electronics 116 can *recognize speech from the user*, convert the speech to AT commands, and control this source device using a standard AT command protocol.”
- [0165]: “Further, the audio device 10A’ can include noise cancellation electronics (not shown) configured to *filter wind-generated noise from an audio signal transmitted from the microphone 75.*”
- [0298]: “In another alternative, the interface electronics 116 can include a speech recognition engine and audio menus. In this alternative, *the interface electronics 116 can recognize speech from the user*, convert the speech to AT commands, and control this source device using a standard AT command protocol.”

123. Jannard teaches “interface electronics 116” “recognize speech from the user.” Jannard, [0298]. As I note above for this limitation in connection with Victorian, a person of ordinary skill in the art would have understood that recognizing speech from the user involves detecting the speech.

124. As I described in §IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface to Victorian’s ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

I. Elements [1.h.4] and [1.h.5]

125. Element [1.h.4] recites “**extracting an audio signal of the user’s voice; and [1.h.5] analyzing the audio signal to determine a voice command.**” These elements are met by Victorian in view of Jannard.

126. Jannard discloses extracting an audio signal of the user’s voice (*e.g.*, “recognize speech from the user”) and analyzing the audio signal to determine a voice command (*e.g.*, “convert the speech to AT commands” or voice commands for the audio device). *See, e.g.*, Jannard:

- [0297]: “For example, if a user chooses a voice command option, the user could issue commands such as, for example, but without limitation, “phonebook” or “call alpha.” With a source device such as a cellular phone, that has a speech recognition engine and that is properly trained to ***recognize the voice of the user***, the user can automatically enter the phonebook mode or directly call the phonebook listing “alpha,” of course, as is apparent to one of ordinary skill in the art, such a voice command protocol could be used to issue other commands as well.”
- [0298]: “In another alternative, the interface electronics 116 can include ***a speech recognition engine and audio menus***. In this alternative, the ***interface electronics 116 can recognize speech from the user***, convert the speech to AT commands, and control this source device using a standard AT command protocol.”

- [0265]: “In a further embodiment, the microphone banks can be arranged in an array to be used in conjunction with algorithms to discern, reduce, and/or eliminate noise for the purpose of voice recognition.”
- [0315]: “The control interface 320 generally includes a buffer, register, pre-processor, transistor, resistor and/or other electronic circuit to ***enable the processor 312 to receive commands from a user via the user controls 322***. In one embodiment, the control interface 320 is integrated with the processor 312.”
- [0214]: “In addition, the selector 610 may be a voice-activated switch, or a voice-activated control that controls selection between primary and secondary sources 602, 604 based upon verbal commands provided by the wearer.”
- [0236]: “The wearer can then elect whether to accept the call, such as by pressing a button on the user interface, ***activating the phone by voice command***, or other initiating mechanism. When the phone call has ended, the user may press a button or activate another control, disconnecting the telephone and resuming listening to data from the primary source.”

127. See my discussion of element [1.h.3] above (§IX.A.4.k).

128. Jannard further teaches “interface electronics 116” “recognize speech from the user” and “convert the speech to AT commands.” Jannard, [0298]. A person of ordinary skill in the art would have understood, and at least found obvious, that recognizing speech from the user such that it can be converted to AT commands involves extracting an audio signal containing the user’s voice from the received audio. Audio signal inputs can include both the user’s speech and ambient noise. The user’s speech is then extracted from the audio signal such that it is separated from the ambient noise so that a voice command can be identified.

129. Jannard further teaches after “recogniz[ing] speech from the user,

convert[ing] the speech to AT commands.” Jannard [0298]. A person of ordinary skill in the art would have understood, and at least found it obvious, that converting the speech to AT commands includes analyzing the audio to determine the voice command. In order for the AT commands to accurately reflect what the voice command is intending, the device would analyze the audio with the voice command.

130. AT commands are part of a protocol for sending commands to a connected device (such as a PDA). Jannard does not separately describe the processing of “voice commands” for operation of the audio device itself. But a person of ordinary skill in the art would have understood (and at least found it obvious) that the processing would be the same (including recognizing speech from the user) as for AT command processing based on voice commands, except that the speech signal would be converted into commands specifically for the audio device itself instead of into AT commands. In both instances, the device takes instructions in the form of recognized speech and maps that recognized speech to a corresponding command.

131. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface to Victorian’s ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice

interface to Victorian, it would be implemented in processor 460.

m. Element [2.pre] and [2.a]

132. Element [2.pre] and [2.a] recite “An earpiece according to claim 1, further including the operations of: analyzing the audio signal to determine a voice command.” See my discussion of claim [1] above (§IX.A.4.a-1). The added limitation of element [2.a] is met by Jannard.

133. This limitation is identical to limitation [1.h.5]. As I explain in connection with limitation [1.h.5], Jannard discloses analyzing the audio signal to determine a voice command (*e.g.*, “recognize speech from the user, convert the speech to AT commands”).

n. Element [2.b]

134. Element [2.b] recites “initiating a response in response to the voice command.” This limitation is met by Jannard.

135. Jannard discloses initiating a response (*e.g.*, transmission of “AT commands” to “source device”; “skip” and audio file, or “delete a playlist”) in response to the voice command (*e.g.*, “convert the speech to AT commands” or voice commands for the audio device). *See, e.g.*, Jannard:

- [0010]: “In general, the data input management system in accordance with the present invention enables the user to select one or more *data sources* to be placed either simultaneously or one at a time into electrical communication with a single user interface...In one configuration, *the user may select one source as a primary input source and a second source as a secondary input source*. One primary input source might

comprise a source of music, such as an MP3 player. A secondary input source might comprise a cellular telephone. *The user may toggle between the input sources, to provide input to the user either automatically, or in response to demand by the user or other electronic prioritization protocol.*

- [0174]: “As illustrated in FIG. 8, the audio device 10C can be worn on the head 18 of a user U. Preferably, the *audio device 10C is configured to provide one or two-way wireless communication with a source device, or the source device can be incorporated into the audio device 10C.* The source device can be carried by the user U, mounted to a moveable object, stationary, or part of a local area or personal area network.”
- [0290]: “According to the headset profile, certain commands can be issued from a headset, such as the audio devices 10, 10A, 10A', 10B, 10C, 10D, and 10E, using an AT command protocol. In such a protocol, text commands must be input to the BLUETOOTH™ device, which the BLUETOOTH™ device then transmits wirelessly to a synchronized BLUETOOTH™ enabled device. *Such commands include, for example, but without limitation, initiating a call, terminating a call, and redialing a previously dialed number.*”
- [0296]: “For example, *when a user chooses the voice command option, the interface electronic 116 can reconfigure to send an AT command to the source device, such as a cellular phone, to accept voice commands directly from the transceiver 114. Thus, as the user speaks, the audio signal is directed to the source device, which in turn is configured to issue audio indicators back to the user, through the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E, to guide the user through such a voice command.*”
- [0297]: “For example, if a user chooses a *voice command option, the user could issue commands such as, for example, but without limitation, “phonebook” or “call alpha.”* With a source device such as a cellular phone, that has a speech recognition engine and that is *properly trained to recognize the voice of the user, the user can automatically enter the phonebook mode or directly call the phonebook listing “alpha,”* of course, as is apparent to one of ordinary skill in the art, *such*

a voice command protocol could be used to issue other commands as well.”

- [0298]: “In another alternative, the interface electronics 116 can include a speech recognition engine and audio menus. In this alternative, *the interface electronics 116 can recognize speech from the user, convert the speech to AT commands, and control this source device using a standard AT command protocol.*”
- [0316]: “*In one embodiment, the user controls 322 include a microphone that receives a voice command. The user controls 322 may be responsive to any biometric provided by a user to control the audio device 304.*”
- [0317]: “In other embodiments, user controls 322 are used to perform any one or a combination of various functions with respect to an audio data file. For example, user controls 322 may be used to *fast-forward, skip, cue, play, pause, turn power on or off, rewind, review, adjust volume, balance, tone, bass, or treble, randomize file selection, load a playlist, set a playlist, delete a playlist, repeat playback of all audio files, selected audio files, or a playlist, or perform any other function related to an audio data file.*”

136. See my discussion of element [2.a] above (§IX.A.4.m).

137. Jannard teaches “the interface electronics 116 can recognize speech from the user, convert the speech to AT commands, and control this source device using a standard AT command protocol.” [0298]. The transmission of an AT command is a “response” to the voice command from which the AT command was converted. Moreover, a person of ordinary skill in the art would have understood that an action performed by the source device—e.g., “initiating a call”—in response to the AT command is initiated by the transmission of the AT command. Jannard [0290]. AT commands are part of a protocol for sending commands to a

connected device (such as a PDA). These commands are used to cause a responsive action, such as initiating a call, when a user speaks a command to initiate that call. Jannard, [0296] (“[C]ommands can be issued using an AT command protocol, to which the source device will also be configured to respond.”).

138. Jannard further teaches initiating other operations (responses) to “voice commands,” including “skip” an audio file and “delete a playlist” on the audio device. Jannard [0316]-[0317].

139. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface to Victorian’s ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

o. Claim [3]

140. **Claim [3] recites “An earpiece according to claim 2, where the voice command is to create an audio content wish list.”** See my discussion of claim [2], above (§IX.A.4.m-n). The added limitation of claim [3] is met by Jannard.

141. Jannard discloses that the voice command (*see* [1.h.5]) is to create an audio content wish list (*e.g.*, using the “voice command option” to provide

“preferred information, which includes preferred song names, artist names, or show or program titles”). *See, e.g., Jannard:*

- [0316]: “In one embodiment, the user controls 322 include a button, dial, switch, lever, sensor, touchpad, microphone, and/or any other input device that may be used by a user to control the audio device 304. ***In one embodiment, the user controls 322 include a microphone that receives a voice command.*** The user controls 322 may be responsive to any biometric provided by a user to control the audio device 304. For example, in one embodiment, the audio device 304 may monitor eye movement, and control the audio device 304 based upon blinking of the user's eyes.”
- [0328]: “The audio device 304 compares the audio signal information to stored preference information to determine whether to receive and store the wireless signal corresponding to a particular song. ***For example, in one embodiment, a user provides the audio device 304 with a list of preferred information, which includes preferred song names, artist names, or show or program titles. The audio device 304 monitors wireless signals to determine if preferred information is or will be broadcast. If the audio device 304 determines that preferred information is or will be broadcast, the audio device 304 receives the preferred information and saves it in the memory module 330.*** In one embodiment, the audio device 304 receives the preferred information and compresses it according to any of the compression or de-compression schemes described in greater detail above. Once compressed, the preferred information is saved in the memory module 330.”
- [329]: “In another embodiment, a user provides the audio device 304 with a list of preferred information, which includes preferred song names, artist names, or show or program titles. The audio device 304 also receives a program guide, which generally indicates the schedule of transmission of songs, programs, or other content from a content source 302, such as the content source 302 shown in FIG. 20. The audio device 304 determines when preferred information will be available by comparing the preferred information to the program guide. The audio device 304 receives and stores the preferred information based upon the comparison.”

142. See my discussion of element [1.h.5] above (§IX.A.4.1).

143. As I explained in connection with element [1.h.5], Jannard teaches user controls 322 allow a user to provide “voice commands.” Jannard, [0298], [0316]. Jannard further teaches a user creates “a list of preferred information, which includes preferred song names, artist names, or show or program titles.” Jannard, [0328]. A person of ordinary skill in the art would have understood (or at least found it obvious) that the list of preferred information is provided through voice commands to advantageously allow the user to provide the preferred information while keeping the user’s hands free to perform other tasks. See my discussion above in Section IX.A.3 regarding the known trend to implement on smaller portable devices using voice controls.

144. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface to identify a list of preferred audio to Victorian’s ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

p. Elements [4.pre]-[4.g]

145. **Element [4.pre] recites “An earpiece.”** This element is met by Victorian. See my discussion of the element [1.pre] above (§IX.A.4.a).

146. **Element [4.a] recites “a sealing section.”** This element is met by Victorian. See my discussion of the element [1.a] above (§IX.A.4.b).

147. **Element [4.b] recites “an ambient microphone.”** This element is met by Victorian. See my discussion of the element [1.b] above (§IX.A.4.c).

148. **Element [4.c] recites “a speaker.”** This element is met by Victorian. See my discussion of the element [1.d] above (§IX.A.4.e).

149. **Element [4.d] recites “an interaction element.”** This element is met by Victorian in view of Jannard. See my discussion of the element [1.e] above (§IX.A.4.f).

150. **Element [4.e] recites “a location unit.”** This element is met by Victorian in view of Jannard. See my discussion of the element [1.f] above (§IX.A.4.g).

151. **Element [4.f] recites “a memory that stores instructions.”** This element is met by Victorian in view of Jannard. See my discussion of the element [1.g] above (§IX.A.4.h).

152. **Element [4.g] recites “a logic circuit configured to execute the instructions to perform operations, the operations comprising.”** This element is met by Victorian in view of Jannard. See my discussion of the element [1.h] above (§IX.A.4.h).

q. Element [4.g.1]

153. Element [4.g.1] recites “detecting the manual activation of the interaction element.” This element is met by Victorian in view of Jannard.

154. Jannard discloses detecting the manual activation of the interaction element (e.g., “user can depress the button 150”). See, e.g., Jannard:

- [0016]: “The selector may comprise *a wearer activated control such as a button*, switch or voice activated electronic control.”
- [0122]: “For example, where the electronic device includes *buttons for controlling an aspect of the device...*”
- [0216]: “[A] *manually activated control such as a button* or touch pad located on an earstem...”
- [0291]: “If the user wishes to initiate a call, the user can depress the button 150 again which will send the appropriate AT command to the transceiver 114 so as to transmit the proper AT code to the cellular phone source device S, B (FIG. 8).”
- [0293]: “[I]f a user wishes to access the phonebook, *the user can depress the button 150 to initiate the audio menu....* Once the user hears the word “phonebook,” the user can *depress the button 150 again to enter the phonebook*. Thereafter, the user can *depress the portions 156, 158 to “scroll” through phonebook entries*. As the user scrolls through the phonebook entries, the interface 116 can be configured to cause the cellular phone to scroll through the phonebook and thereby transmit an audio signal of a humanoid voice indicating entries in the phonebook. When the user hears the name of the person or entity which the user desires to call, the user can again *push the button 150 to initiate a call* to that entity.”

155. See my discussion for element [1.e], above (§IX.A.4.f).

156. Jannard further teaches scrolling through audio menu options until a “voice command option” is identified, at which point “the user can depress the button 150” to initiate the use of voice commands. Jannard, [0296].

157. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard's teachings of a voice interface activated by a button to Victorian's ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard's teachings of a voice interface to Victorian, it would be implemented in processor 460.

r. Element [4.g.2]

158. **Element [4.g.2] recites “switching the mode of the earpiece to voice command mode when manual activation of the interaction element is detected.”** This element is met by Victorian in view of Jannard.

159. **Jannard discloses switching the mode of the earpiece to voice command mode** (*e.g.*, “chooses the voice command option”) **when manual activation of the interaction element is detected** (*e.g.*, “depress the button 150”).

See, e.g., Jannard:

- [0292] (“*[T]he user can depress either the forward or rearward portions 156, 158 of the button 150 so as to “scroll” through other audio menu options. For example, other audio menu options can include, for example, but without limitation, phonebook, email, clock, voice commands, and other menu options typically available on cellular phones and/or personal audio devices such as MP3 players.*”)
- [293] (“As an illustrative, but non-limiting example, *if a user wishes to access the phonebook, the user can depress the button 150 to initiate the audio menu, then “scroll”* to the phonebook by depressing the portions 156 or 158 until the user hears the word “phonebook” in the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E.

Once the user hears the word “phonebook,” the user can depress the button 150 again to enter the phonebook.”)

- [0296]: “Once the user hears the word “reply” the user can depress the button 150 to enter a reply mode. As noted above, these types of commands can be issued using an AT command protocol, to which the source device will also be configured to respond. As noted above, one audio menu option can include voice command. As noted above, *one audio menu option can include voice command*. For example, *when a user chooses the voice command option*, the interface electronic 116 can reconfigure to send an AT command to the source device, such as a cellular phone, to accept voice commands directly from the transceiver 114.”
- [0298]: “In another alternative, *the interface electronics 116 can include a speech recognition engine* and audio menus. In this alternative, the *interface electronics 116 can recognize speech from the user*, convert the speech to AT commands, and control this source device using a standard AT command protocol.”
- [0316]: “In one embodiment, the user controls 322 include a *microphone that receives a voice command*.”

160. See my discussion for elements [1.h.2], above (§IX.A.4.j).

161. Jannard further teaches a user “chooses the voice command option” by “depress[ing] the button 150.” Jannard, [0296]. This allows the device to switch to operate as an audio interface to receive voice commands. Jannard, [0296].

162. As I described in § IX.A.3, a POSITA would have been motivated to apply Jannard’s teachings of a voice interface activated by a button to Victorian’s ear-level device. As explained in [1.h.2], a POSITA would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be

implemented in processor 460.

s. Elements [4.g.3]-[4.g.6]

163. **Element [4.g.3] recites “detecting a user's voice.”** This element is met by Victorian in view of Jannard. See my discussion of element [1.h.3] above. (§IX.A.4.k).

164. **Element [4.g.4] recites “extracting an audio signal of the user's voice.”** This element is met by Victorian in view of Victorian in view of Jannard. See my discussion of element [1.h.4] above. (§IX.A.4.l).

165. **Element [4.g.5] recites “analyzing the audio signal to determine a voice command.”** This element is met by Victorian in view of Jannard. See my discussion of element [1.h.5] above. (§IX.A.4.l).

166. **Element [4.g.6] recites “initiating a response in response to the voice command.”** This element is met by Victorian in view of Jannard. See my discussion of element [2.b] above. (§IX.A.4.n).

t. Element [4.h]

167. **Element [4.h] recites “where the voice command is perform at least one of the following actions with respect to an audio content list stored in the memory, purchase a song in the audio content list, delete a song from the audio content list, skip to the next song in the audio content list, add a song to the audio content list, and delete the audio content list.”** This element is met by

Victorian in view of Jannard.

168. Jannard discloses the voice command is perform at least one of the following actions with respect to an audio content list stored in the memory: skip to the next song in the audio content list (e.g., “skip”), and delete the audio content list (e.g., “delete a playlist”). See, e.g., Jannard:

- [315]: “The audio device 304 may also include ***a memory module 330, which in one embodiment stores audio data files***. The memory module 330 may include any of a variety of electronic memory devices, including but not limited to, a hard drive, flash memory, RAM, ROM, EPROM, EEPROM, or PROM. In one embodiment, the memory module 330 includes NAND flash memory. In one embodiment, the memory module 330 includes at least about 128 MB. In other embodiments, the memory module 330 includes, by way of example, at least about 256 MB, 512 MB, or 1 GB of memory.”
- [0316]: “In one embodiment, the user controls 322 include a button, dial, switch, lever, sensor, touchpad, microphone, and/or any other input device that may be used by a user to control the audio device 304. ***In one embodiment, the user controls 322 include a microphone that receives a voice command***. The user controls 322 may be responsive to any biometric provided by a user to control the audio device 304. For example, in one embodiment, the audio device 304 may monitor eye movement, and control the audio device 304 based upon blinking of the user's eyes.”
- [0317]: “In other embodiments, user controls 322 are used to perform any one or a combination of various functions with respect to an audio data file. For example, user controls 322 may be used to ***fast-forward, skip, cue, play, pause, turn power on or off, rewind, review, adjust volume, balance, tone, bass, or treble, randomize file selection, load a playlist, set a playlist, delete a playlist, repeat playback of all audio files, selected audio files, or a playlist, or perform any other function related to an audio data file***.”

169. See my discussion for elements [1.h.1]-[1.h.5] above (§IX.A.4.i-1).

170. Jannard further teaches the microphone receives “voice commands” including to perform functions such as “skip” an audio file and “delete a playlist.” Jannard [0316]-[0317].

171. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface used to control media content to Victorian’s ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

u. Claim [5]

172. **Claim [5] recites “An earpiece according to claim 2, where the voice command is to search the internet.”** See my discussion of claim [2] above (§IX.A.4.m-n). The added limitation of claim [6] is met by Jannard.

173. Jannard discloses the voice command (*see* [1.h.5]) is to search the internet (*e.g.*, “internet browsing”). *See, e.g.*, Jannard:

- [0176]: “Optionally, the device B can also be configured to communicate, via long or short range wireless networking protocols, with a remote source R. ***The remote source R can be, for example, but without limitation, a cellular phone service provider, a satellite radio provider, or a wireless internet service provider.*** For example, but without limitation, the source device B can be configured to communicate with other wireless data networks such as via, for example, but without limitation, long-range packet-switched network protocols including PCS, GSM, and GPRS. As such, the audio device 10C can be used as an audio interface for the source device B. For example, but

without limitation, where the source device B is a cellular phone, the user U can listen to the audio output of the cellular phone, such as the voice of a caller, through sound transducers in the audio device 10C. Optionally, the user U can send voice signals or commands to the cellular phone by speaking into a microphone on the audio device 10C, described in greater detail below. Thus, the audio device 10C may advantageously be a receiver and/or a transmitter for telecommunications.”

- [0297]: “For example, if a user chooses a voice command option, the user could issue commands such as, for example, but without limitation, “phonebook” or “call alpha.” With a source device such as a cellular phone, that has a *speech recognition engine* and that is properly trained to recognize the voice of the user, the user can automatically enter the phonebook mode or directly call the phonebook listing “alpha,” of course, as is apparent to one of ordinary skill in the art, such *a voice command protocol could be used to issue other commands as well.*”
- [0298]: “In another alternative, the interface electronics 116 can include a *speech recognition engine and audio menus*. In this alternative, the interface electronics 116 can recognize speech from the user, convert the speech to AT commands, and control this source device using a standard AT command protocol.”
- [0299]: “*For example, but without limitation, the source device B can be in the form of a palm-top or hand-held computer* known as a BLACKBERRY™. The presently marketed BLACKBERRY™ devices can communicate with a variety of wireless networks for receiving email, phone calls, and/or *internet browsing*. One aspect of at least one of the present inventions includes the realization that such a hand-held computer can include a text-to-speech engine. Thus, such a hand-held computer can be used in conjunction with any of the audio devices 10, 10A, 10A', 10B to allow a user to “hear” emails, or other text documents without the need to hold or look at the device B. Preferably, the hand-held computer includes a further wireless transceiver compatible with at least one of the transceivers 114, 114 i. As such, a user can use any of the audio devices 10C, 10D, 10E to “hear” emails, or other text documents without the need to hold or look at the device B. Thus, a presently preferred hand-held computer, as a non-limiting example, includes a BLACKBERRY™ hand-held device including long range wireless network hardware for email and internet browsing capability, a

BLUETOOTH™ transceiver for two-way short range audio and/or data audio communication, and a text-to-speech engine.”

- [0301]: “An audio network 300 in accordance with another embodiment of the present invention is illustrated in FIG. 20. Audio network 300 includes a content source 302 coupled to an audio device 304 via communications link 306. The content source 302 is any of a variety of information sources, including, but not limited to, radio stations and/or signals, a satellite radio source, a computer, a network, a storage device, such as a hard drive, a memory card, or a USB (Universal Serial Bus) drive, an audio component (e.g., a stereo receiver, a CD player, a tuner, an MP3 player, a digital audio player, etc.), a database, and/or a communications-enabled device, such as a telephone (including a BLUETOOTH enabled telephone), a PDA, a BLACKBERRY, *the Internet, or the like*. The content provided by the content source 302 may be any of a variety of information, including but not limited to, audio files, entertainment, news, media, music, photos, videos, advertising, etc.”

174. See my discussion for elements [1.h.5] above (§IX.A.4.1).

175. Jannard further teaches “internet browsing” in response to “recogniz[ing] speech from the user, convert[ing] [] speech to AT commands, and control[ing]” the “source device using a standard AT command protocol.”

Jannard, [0298]. A person of ordinary skill in the art would have understood that “internet browsing” includes the use of a search engine to perform a web search to allow a user to find a desired website, and that such a web search would provide resulting search results. At the very least, a person of ordinary skill in the art would have found it obvious to use a search engine to find desired content from the large number of websites available on the internet. *See, e.g.,* Ex. 1013 (Martin), [0005] (“Because of the large number of websites on the Internet, searching the

Internet for specific information usually requires the services of a search engine. Generally speaking, a search engine helps consumers look for relevant web pages...The links to matching web pages are presented to the consumer in the form of a listing referred to as “search results.””). Thus, internet browsing is understood by a person of ordinary skill in the art to include web searching, or at the very least, a user would perform web searching. Even more, Jannard discloses using “text-to-speech” for “internet browsing,” such that when the user performs a web search, the search results would be provided to the user via text-to-speech. Jannard, [0299].

176. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface for internet browsing to Victorian’s ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

v. Claim [6]

177. **Claim [6] recites “An earpiece according to claim 5, where the internet search results are received via a text to speech analyzer resulting in an audio result and the user receives the audio result.”** See my discussion of claim [5] above (§IX.A.4.u). The added limitation of claim [6] is met by Jannard.

178. Jannard discloses where the internet search results (*see* [5]) are received via a text to speech analyzer (*e.g.*, “text-to-speech engine”) resulting in an audio result and the user receives the audio result (*e.g.*, a user can “‘hear’ emails or other text documents” received from the internet). *See e.g.*, Jannard:

- [0295]: “***A text-to-speech engine can provide further convenient uses for a user.*** For example, if the cell phone or other source device is configured to receive email, the device can be configured to signal the user with an audio signal that an email has been received. The user can send a signal to the source device so as to open the email. ***The text-to-speech engine can be configured to read the email to the user.*** Thus, a user can “listen” to email through the audio device 10, 10A, 10A', 10B, 10C, 10D, 10E, without manually operating the source device.”
- [0296]: “A further option is to allow a user to respond to such an email. For example, the user could record an audio file, such as, for example, but without limitation a .WAV, .MP3 file as an attachment to a reply email. For such a feature, the interface 116 can be configured to automatically provide a user with options at the end of an email that is read to the user. ***For example, after the text-to-speech engine finishes “reading” the email to the user, the interface device 116 can enter another audio menu.*** Such an audio menu can include a reply option, a forward option, or other options. If a user wishes to reply, the user can “scroll” until the user hears the word “reply.” Once the user hears the word “reply” the user can depress the button 150 to enter a reply mode. As noted above, these types of commands can be issued using an AT command protocol, to which the source device will also be configured to respond. As noted above, one audio menu option can include voice command. For example, when a user chooses the voice command option, the interface electronic 116 can reconfigure to send an AT command to the source device, such as a cellular phone, to accept voice commands directly from the transceiver 114. Thus, as the user speaks, the audio signal is directed to the source device, which in turn is configured to issue audio indicators back to the user, through the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E, to guide the user through such a voice command.”

- [0299]: “For example, but without limitation, the source device B can be in the form of a palm-top or hand-held computer known as a BLACKBERRY™. The presently marketed BLACKBERRY™ devices can communicate with a variety of wireless networks for receiving email, phone calls, and/or *internet browsing*. One aspect of at least one of the present inventions includes the realization that ***such a hand-held computer can include a text-to-speech engine***. Thus, such a hand-held computer can be used in conjunction with any of the audio devices 10, 10A, 10A', 10B to ***allow a user to “hear” emails, or other text documents without the need to hold or look at the device B. Preferably, the hand-held computer includes a further wireless transceiver compatible with at least one of the transceivers 114, 114 i. As such, a user can use any of the audio devices 10C, 10D, 10E to “hear” emails, or other text documents without the need to hold or look at the device B.*** Thus, a presently preferred hand-held computer, as a non-limiting example, includes a BLACKBERRY™ hand-held device including long range wireless network hardware for email and internet browsing capability, a BLUETOOTH™ transceiver for two-way short range audio and/or data audio communication, and a text-to-speech engine.”
- [194]: “The source electronics may be a microphone or a radio, such as a terrestrial-based or satellite-based radio, including XM® or SIRIUS® brand satellite radios. In other embodiments, the source electronics may be a telephone, a cellular telephone, a personal digital assistant (PDA), a BLACKBERRY®, or a pager. A variety of currently available devices, for example, a BLACKBERRY®, pager, any of a variety of PDA's and e-mail enabled devices such as notebook computers provide incoming text messages. In one aspect of the present invention any of these text message devices is provided with text to voice software, enabling the text to be read out loud. This enables the user to listen to a primary source such as music, or the sound track from a portable DVD player, and incoming e-mails will be read out loud to the wearer, while the primary source is either placed on pause, or remains running in the background. Text to voice software can either be carried by the support, or carried by the underlying source such as the BLACKBERRY® or other PDA.”

179. See my discussion for elements [1.h.5] and [5] above (§IX.A.4.1, § IX.A.4.u).

180. Jannard further discloses using “text-to-speech” for “internet browsing,” such that when the user performs a web search, the search results would be provided to the user via text-to-speech. Jannard, [0299].

181. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface for internet browsing to Victorian’s ear-level device. A person of ordinary skill in the art further would have also been motivated to apply Jannard’s teachings of text-to-speech playback of internet pages so as to advantageously allow the user to browse the internet in a hands-free manner, as I explain above in §IX.A.3. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

w. Claim [7]

182. **Claim [7] recites “An earpiece according to claim 2, where the voice command is to play audio from the internet.”** See my discussion of the elements of claim [2] above (§IX.A.4.m-n). The added limitation of claim [7] is met by Jannard.

183. Jannard discloses the voice command (*See* [1.h.5]) is to play audio from the internet (*e.g.*, “play” “audio files” from “the Internet”). *See, e.g.*, Jannard:

- [292]: “The user can be provided with further convenience if there are other menu choices available, for example, if the user does not wish to

- choose the first menu option, *the user can depress either the forward or rearward portions 156, 158 of the button 150 so as to “scroll” through other audio menu options. For example, other audio menu options can include*, for example, but without limitation, phonebook, email, clock, *voice commands*, and other menu options typically available on cellular phones and/or personal audio devices such as MP3 players.”
- [0296]: “A further option is to allow a user to respond to such an email. For example, the user could record an audio file, such as, for example, but without limitation a .WAV, .MP3 file as an attachment to a reply email. For such a feature, the interface 116 can be configured to automatically provide a user with options at the end of an email that is read to the user. For example, after the text-to-speech engine finishes “reading” the email to the user, *the interface device 116 can enter another audio menu. Such an audio menu can include a reply option, a forward option, or other options. If a user wishes to reply, the user can “scroll” until the user hears the word “reply.” Once the user hears the word “reply” the user can depress the button 150 to enter a reply mode.* As noted above, these types of commands can be issued using an AT command protocol, to which the source device will also be configured to respond. *As noted above, one audio menu option can include voice command.* For example, *when a user chooses the voice command option*, the interface electronic 116 can reconfigure to send an AT command to the source device, such as a cellular phone, *to accept voice commands* directly from the transceiver 114. Thus, as the user speaks, the audio signal is directed to the source device, which in turn is configured to issue audio indicators back to the user, through the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E, to guide the user through such a voice command.”
 - [0297]: “For example, if a user chooses a voice command option, the user could issue commands such as, for example, but without limitation, “phonebook” or “call alpha.” With a source device such as a cellular phone, that has a speech recognition engine and that is properly trained to recognize the voice of the user, the user can automatically enter the phonebook mode or directly call the phonebook listing “alpha,” of course, as is apparent to one of ordinary skill in the art, such *a voice command protocol could be used to issue other commands as well.*”
 - [0298]: “In another alternative, the interface electronics 116 can include a

speech recognition engine and audio menus. In this alternative, the interface electronics 116 can *recognize speech from the user*, convert the speech to AT commands, and control this source device using a standard AT command protocol.”

- [301]: “An audio network 300 in accordance with another embodiment of the present invention is illustrated in FIG. 20. Audio network 300 includes *a content source 302 coupled to an audio device 304 via communications link 306. The content source 302 is any of* a variety of information sources, including, but not limited to, radio stations and/or signals, a satellite radio source, a computer, a network, a storage device, such as a hard drive, a memory card, or a USB (Universal Serial Bus) drive, an audio component (e.g., a stereo receiver, a CD player, a tuner, an MP3 player, a digital audio player, etc.), a database, and/or a communications-enabled device, such as a telephone (including a BLUETOOTH enabled telephone), a PDA, a BLACKBERRY, *the Internet*, or the like. *The content provided by the content source 302 may be any of a variety of information, including but not limited to, audio files*, entertainment, news, media, music, photos, videos, advertising, etc.”
- [0316]: “In one embodiment, the user controls 322 include a button, dial, switch, lever, sensor, touchpad, microphone, and/or any other input device that may be used by a user to control the audio device 304. *In one embodiment, the user controls 322 include a microphone that receives a voice command.* The user controls 322 may be responsive to any biometric provided by a user to control the audio device 304. For example, in one embodiment, the audio device 304 may monitor eye movement, and control the audio device 304 based upon blinking of the user's eyes.”
- [0317]: “In other embodiments, user controls 322 are used to perform any one or a combination of various functions with respect to an audio data file. For example, user controls 322 may be used to fast-forward, skip, cue, *play*, pause, turn power on or off, rewind, review, adjust volume, balance, tone, bass, or treble, randomize file selection, load a playlist, set a playlist, delete a playlist, repeat playback of all audio files, selected audio files, or a playlist, or perform any other function *related to an audio data file.*”

184. See my discussion of elements [1.h.5] and claims [5]-[6] above

(§IX.A.4.1, § IX.A.4.u-v).

185. Jannard further teaches “user controls 322,” which include a microphone to receive voice commands “used to...play” an “audio data file.” Jannard, [0316], [0317]. Jannard further teaches “audio files” are retrieved from “the Internet.” Jannard, [0301]. For example, a person of ordinary skill in the art would understand the audio device connected to the internet is capable of downloading files from a webpage, from an FTP site, or via an email attachment. A person of ordinary skill in the art thus would have understood (or at least found it obvious) that audio files from the internet (either downloaded or streamed) are played in response to voice commands. In this regard, as I explain in Section IX.A.3, by 2006, it was recognized that “[m]any wireless cellular phones now provide “wireless web” limited browsing capability,” and that “use of these products supports the trend of access to Internet content beyond the PC.” Ex. 1019 (Janik), [0018].

186. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface for media content control and internet browsing to Victorian’s ear-level device. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

x. Claim [9]

187. Claim [9] recites “An earpiece according to claim 2, where the voice command is to play audio from a radio station.” See my discussion of the elements of claim [2] above (§IX.A.4.m-n). The added limitation of claim [9] is met by Jannard.

188. Jannard discloses the voice command (*see* [1.h.5]) is to play audio from a radio station (*e.g.*, “play” “audio” from “FM or satellite radio”). *See, e.g.*, Jannard:

- [0176]: “Optionally, the device B can also be configured to communicate, via long or short range wireless networking protocols, with a remote source R. ***The remote source R can be, for example, but without limitation, a cellular phone service provider, a satellite radio provider, or a wireless internet service provider.*** For example, but without limitation, the source device B can be configured to communicate with other wireless data networks such as via, for example, but without limitation, long-range packet-switched network protocols including PCS, GSM, and GPRS. As such, the audio device 10C can be used as an audio interface for the source device B. For example, but without limitation, where the source device B is a cellular phone, the user U can listen to the audio output of the cellular phone, such as the voice of a caller, through sound transducers in the audio device 10C. Optionally, the user U can send voice signals or commands to the cellular phone by speaking into a microphone on the audio device 10C, described in greater detail below. Thus, the audio device 10C may advantageously be a receiver and/or a transmitter for telecommunications.”
- [194]: “The source electronics may be a microphone or a radio, such as a terrestrial-based or satellite-based radio, including XM® or SIRIUS® brand satellite radios. In other embodiments, the source electronics may be a telephone, a cellular telephone, a personal digital assistant (PDA), a BLACKBERRY®, or a pager. A variety of currently available devices, for example, a BLACKBERRY®, pager, any of a variety of PDA's and e-

mail enabled devices such as notebook computers provide incoming text messages. In one aspect of the present invention any of these text message devices is provided with text to voice software, enabling the text to be read out loud. This enables the user to listen to a primary source such as music, or the sound track from a portable DVD player, and incoming e-mails will be read out loud to the wearer, while the primary source is either placed on pause, or remains running in the background. Text to voice software can either be carried by the support, or carried by the underlying source such as the BLACKBERRY® or other PDA.”

- [0298]: “In another alternative, the interface electronics 116 can include a *speech recognition engine* and audio menus. In this alternative, the interface electronics 116 can *recognize speech from the user*, convert the speech to AT commands, and control this source device using a standard AT command protocol.”
- [0301]: “An audio network 300 in accordance with another embodiment of the present invention is illustrated in FIG. 20. Audio network 300 includes a content source 302 coupled to an audio device 304 via communications link 306. The content source 302 is any of a variety of information sources, including, but not limited to, *radio stations* and/or signals, a satellite radio source, a computer, a network, a storage device, such as a hard drive, a memory card, or a USB (Universal Serial Bus) drive, an audio component (e.g., a stereo receiver, a CD player, a tuner, an MP3 player, a digital audio player, etc.), a database, and/or a communications-enabled device, such as a telephone (including a BLUETOOTH enabled telephone), a PDA, a BLACKBERRY, the Internet, or the like. The content provided by the content source 302 may be any of a variety of information, including but not limited to, audio files, entertainment, news, media, music, photos, videos, advertising, etc.”
- [0316]: “In one embodiment, the user controls 322 include a button, dial, switch, lever, sensor, touchpad, microphone, and/or any other input device that may be used by a user to control the audio device 304. In one embodiment, *the user controls 322 include a microphone that receives a voice command*. The user controls 322 may be responsive to any biometric provided by a user to control the audio device 304. For example, in one embodiment, the audio device 304 may monitor eye movement, and control the audio device 304 based upon blinking of the

user's eyes.”

- [0317]: “In other embodiments, user controls 322 ***are used to perform any one or a combination of various functions with respect to an audio data file***. For example, user controls 322 may be used to fast-forward, skip, cue, ***play***, pause, turn power on or off, rewind, review, adjust volume, balance, tone, bass, or treble, randomize file selection, load a playlist, set a playlist, delete a playlist, repeat playback of all audio files, selected audio files, or a playlist, or perform any other function related to an audio data file.”
- [0327]: “In one embodiment, the audio device 304 receives a wireless signal, such as an *FM or satellite radio*, or wireless network, infrared, Bluetooth signal, or the like. The audio device 304 identifies audio signal information in the wireless signal, such as the performer of a song corresponding to the wireless signal.”

189. See my discussion of claim [7] above (§IX.A.w).

190. As I explain in connection with claim [7], Jannard teaches the user may provide “voice commands” to play received audio files. Jannard further teaches storing audio received via “a wireless signal, such as an FM or satellite radio.” Jannard, [0327]. A person of ordinary skill in the art thus would have understood (or at least found it obvious) that audio files from FM or satellite radio are played in response to voice commands.

191. As I described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Jannard’s teachings of a voice interface to Victorian’s ear-level device. Moreover, a person of ordinary skill in the art would have been motivated to use voice controls to start playing audio from the radio station to advantageously provide hands-free control. For example, as I explain in

Section IX.A.3, by 2006, it was recognized that “[m]any wireless cellular phones now provide “wireless web” limited browsing capability,” and that “use of these products supports the trend of access to Internet content beyond the PC.” Ex. 1019 (Janik), [0018]. As I described in [1.h.2], a person of ordinary skill in the art would have understood that, when applying Jannard’s teachings of a voice interface to Victorian, it would be implemented in processor 460.

y. Claim [15]

192. **Claim [15] recites “The earpiece according to claim 2 including an ear canal microphone.”** See my discussion of the elements of claim [2] above (§IX.A.4.m-n). The added limitation of this claim is met by Victorian in view of Jannard. See my discussion of element [1.c] above (§IX.A.4.d).

z. Claim [16]

193. **Claim [16] recites “An earpiece according to claim 15, further including the operations of: measuring the sound pressure level using the ear canal microphone.”** See my discussion of the elements of claim [15] above (§IX.A.4.y). The added limitation of claim [16] is met by Victorian.

194. Victorian discloses further including the operations of (*see* [1.g]-[1.h]): measuring the sound pressure level using the ear canal microphone (*e.g.*, “sound detected by ear canal microphone 440 exceeds [a] predetermined threshold”). *See, e.g.*, Jannard:

- [0016] (“These systems each include *one or two ear-level devices* each worn in or about an ear, such as in-the-ear (ITE) or behind-the-ear (BTE) devices, *wirelessly coupled to a remote device such as* a computer, a personal digital assistant (PDA), *a cellular phone*, a walkie talkie, or a language translator.”)
- [0032] (“Ear-level device 110 includes an external microphone 420, an ear canal microphone 440, an ear canal speaker 470, a processor 460, a wireless transceiver 450, an antenna 455, a voice operated exchange (VOX) circuit 430, and a battery 480. External microphone 420 picks up ambient sound such that a hearing impaired person wearing ear-level device 110 is not "isolated" when communicating to remote device 190. Thus, ear-level device enables the hearing impaired person to talk to another person either directly (if the other person is nearby) or through a device such as a cell phone. Ear canal microphone 440 detects sound from the ear canal. The sound includes primarily speech of the person wearing ear-level device 110. Ear canal speaker 470 transmits sound received from remote device 190 and/or external microphone 420 to the same ear canal. In one embodiment, ear canal microphone 440 and ear canal speaker 470 are implemented as one physical device. Processor 460 converts the sound picked up from the ear canal to an electrical signal to be transmitted to remote device 190, and converts the signals received from remote device 190 and/or external microphone 420 to a sound audible to the person wearing ear-level device 110. In one embodiment, processor 460 includes one or more of amplification circuitry, filtering circuitry, acoustic feedback reduction circuitry, noise reduction circuitry, and tone control circuitry, among other circuits performing signal processing functions as known in the art. In one embodiment, processor 460 includes a speech recognition module to enhance the audio signal received by and/or transmitted from ear-level device 110. Wireless transceiver 450 and antenna 455 form a telemetry interface supporting telemetry link 195 between ear-level device 110 and remote device 190. Wireless transceiver 450 includes a wireless transmitter and a wireless receiver. The wireless transmitter receives the electrical signal representing the sound picked up from the ear canal from processor 460 and transmits an outgoing wireless signal representing the same sound to remote device 190. The wireless receiver receives an incoming wireless signal representing a remote sound from remote device 190 and transmits an electrical signal representing the remote sound to processor 460. Telemetry link 195 provides for bi-directional communication allowing

simultaneous signal transmission, in both directions, between ear-level device 110 and remote device 190. VOX circuit 430 activates a major portion of ear-level device 110 only when sound is detected by at least one of external microphone 420, ear canal microphone 440, and wireless transceiver 450. In one embodiment, VOX circuit 430 includes a voice-controlled switch connecting between battery 480 and the portions of ear-level device 110 that are voice activated. The switch is driven by the sound detected by at least one of external microphone 420, *ear canal microphone 440*, and wireless transceiver 450. It is turned on when sound detected by external microphone 420 exceeds a predetermined threshold, *when sound detected by ear canal microphone 440 exceeds another predetermined threshold*, and/or when wireless transceiver 450 receives the incoming wireless signal representing the remote sound. In one embodiment, as discussed with respect to FIG. 4B below, VOX circuit 430 gates or attenuates one or more sounds detected by external microphone 420, ear canal microphone 440, and wireless transceiver 450, to eliminate or reduce echo and ringing caused by the loop formed by ear canal speaker 470 and ear canal microphone 440. Battery 480 supplies the power needed for the operation of ear-level device 110. In one embodiment, battery 480 is a rechargeable battery that can be recharged without being taken out from ear-level device 110.”)

195. See my discussion of elements [1.g]-[1.h] above (§IX.A.4.h).

196. Victorian discloses “VOX circuit 430 activates a major portion of ear-level device 110 only when” “sound detected by ear canal microphone 440 exceeds [a] predetermined threshold.” Victorian, [0030].

197. A person of ordinary skill in the art would have understood that the “sound pressure level” indicates the sound level or sound amplitude. Ex. 1015, Pages 21-22. For example, it was understood that “it is convenient to measure sound amplitude on a logarithmic scale in decibels (dB),” and “[s]ound pressure level (SPL) is a measure of absolute sound pressure P in dB.” SLP, Pages 21-22.

In this regard, Victorian's "predetermined threshold" refers to a level of sound. [0030]. Above that level of sound, a switch turns on and activates the ear-level device. Thus, the sound pressure level is measured to make the comparison to Victorian's "predetermined threshold."

198. Moreover, **Victorian** discloses processor 460 "converts" sound picked up by ear canal microphone 440 for transmission to a remote device, such as a cellular phone. Victorian, [0016], [0030]. A POSITA would have understood (and at least found it obvious) that, as part of this process, the relative sound pressure level picked up by ear canal microphone 440 is determined to maintain the relative sound level loudness for transmission to the remote device, thereby ensuring the sound heard by the user of the remote device is accurate. For example, the microphone picks up sound and generates an analog electrical signal directly proportional to air pressure variation due to the sound. That analog electrical signal then goes into an analog to digital converter that converts it into a time sequence of binary 0's/1's in a particular format on a processor. To do that conversion, there is a "dynamic range" of the analog to digital converter, e.g., +/- 1 volt, and the binary code is representative of where in the range the electrical signal falls. This binary signal maintains the same relative loudness profile, thereby ensure relative loudness based on the measured relative sound pressure level. Such determination of the relative sound pressure level is consistent with the

'632, in which the only disclosure of measuring sound pressure level is that variations in sound pressure level are determined, but the actual sound pressure level itself is not determine. '632, 5:25-29 (“Acoustic Transducer...converts sound pressure level variations into electronic voltages”).

B. GROUND 2: Ground 1 in further view of Karaoguz (Claim 8):

1. Overview of Karaoguz

199. Karaoguz discloses a main device, such as a PC, and a “media peripheral device” such as an “MP3 player” that plays “audio files.” Karaoguz, [0050], [0054]. The PC and MP3 player are connected over a “communication network,” such as “the internet.” Karaoguz, [0024], [0047]. Karaoguz further discloses various “commands” for manipulating the media, including “scan.” Karaoguz, [0024], [0052].

200. A person of ordinary skill in the art would have understood that the general reference to “scan” in Karaoguz is consistent with the use of “scan” in the '632 patent. In particular, a person of ordinary skill in the art would have understood that, to “scan audio from the internet” is consistent with the existing scan of audio from, e.g., a radio. In particular, the '632 patent describes “playback of short audio clips” as “similar to the ‘scan’ button interface common in FM/AM radio devices.” '632 patent, 11:53-63. In other words, to “scan audio from the internet” means to play short clips of audio from the internet, consistent with the

common usage of “scan” in this field. For example, “[t]he seek and scan functions of a radio sequentially tune to the next consecutive strong signal. The seek function must be continuously engaged by a listener to set to the next sequentially strong signal, whereas the scan function continues to search for the next sequentially strong signal but incorporates a delay for the listener to hold a signal thereby discontinuing a search of the next strong signal. Additionally, a listener may manually program a limited number of station presets and may be able to seek and scan through them as well.” *Srail*, Ex. 1006 [0002].

2. Motivation to Apply the Karaoguz’s Teachings to Victorian and Jannard

201. Victorian, Jannard, and Karaoguz are analogous art. Victorian, Jannard, and Karaoguz are in the same field as the ’632 patent, which includes personal audio devices ’632, Abstract, 3:57-4:3; Victorian, [0016], [0049]; Jannard, [0276], [0292]; Karaoguz, [0007], [0016], [0017]; see §IX.A.3. Victorian, Jannard, and Karaoguz are reasonably pertinent to alleged problems addressed by the ’632 patent of customizing and controlling media selection in a personal device. *E.g.*, ’632 patent, 3:39-41, (“The system is designed to operate as simply as possible. Using a single button, which has multiple contacts, the interface allows the user to purchase, delete, skip to next, and add to a wish list and even control a listening level.”); Victorian, [0049] (“The method provides for private, simultaneous two-way communication between a person and a remote

device such as a computer, a personal digital assistant (PDA), a cellular phone, a walkie talkie, or a language translator.”); Jannard, [0235] (“Thus, in one embodiment the user selects a music source as the primary source 602, and may select a cellular telephone as the secondary source 604, with an override capability to enable interruption of the music to reveal the source of the incoming phone call.”); Karaoguz, [0016] (“[T]he present invention relate[s] to indirectly monitoring and controlling media peripheral devices on a media exchange network.”).

202. A person of ordinary skill in the art would have been motivated to apply Karaoguz’s teachings of executing a “command” to scan “audio” to Victorian’s “ear-level device” (as modified by Jannard). For example, Jannard describes using voice commands to obtain content from an “audio network,” which includes a “content source” such as “the Internet,” but does not enumerate that this particular feature can be implemented with a scanning function for the user to preview a set of audio content options. Jannard, [0297], [0298], [0301], [0303], [0327], [0328], [0332]. A person of ordinary skill in the art would have been motivated to look to analogous art—such as Karaoguz—to advantageously allow the user to preview samples of audio content by scanning through the samples of that content so that the user can make a more informed decision about a desirable audio content to select and download. *See* Karaoguz, [0018], [0021], [0052]. For

example, it was recognized at the time that internet connections on small device were slow due to the “slow data-rates provided by these products.” Janik, [0018], Moreover, extending the audio selection capabilities to include audio scanning as taught by Karaoguz would be a simple and straightforward modification.

203. Thus, a person of ordinary skill in the art would have found it routine, straightforward and advantageous to apply Karaoguz’s teachings of scanning audio by voice command to Victorian’s ear-level device (as modified by Jannard’s teachings of a voice control interface) and would have known such a combination (yielding the claimed limitations) would predictably work and provide the expected functionality.

3. Claim [8]

204. **Claim [8] recites “An earpiece according to claim [7], where a second voice command is to scan audio from the internet.”** See my discussion of claim [7], above (§IX.A.4.w). The limitation added by this claim is met by Victorian in view of Jannard and Karaoguz.

205. As I explain in connection with [1.h.1]-[1.h.5] and [4.h] (§§IX.A.4.i-l) Jannard discloses voice commands provided via user controls 322, including commands to control audio content.

206. Karaoguz discloses voice commands to scan audio from the internet (a “voice...remote control” used to “scan” “audio files” from a “communication

network”). See, e.g., Karaoguz:

- [0024]: “In an embodiment of the present invention, ***the media may comprise at least one of audio***, a still image, video, real-time video, and data. The at least one media peripheral command may comprise at least one of on, off, select, play, capture, download, erase, delete, zoom, focus, pan, tilt, set compression format, set resolution, set frame rate, set quality, rewind, fast forward, ***scan***, list, skip, and check status. ***The communication network may comprise at least one of a cable infrastructure, a satellite network infrastructure, a digital subscriber line (DSL) infrastructure, an Internet infrastructure, an intranet infrastructure, a wired infrastructure, and a wireless infrastructure.*** The at least one media peripheral may comprise at least one of a digital camera, a digital camcorder, an MP3 player, a home juke-box system, a multi-media personal digital assistant (PDA), a mobile multi-media gateway device.”
- [0047]: “The WAN(s) 103 may include cable infrastructure, DSL infrastructure, ***Internet infrastructure***, or intranet infrastructure in order to provide communications between, for example, the home location 115, the remote location 116, and third party locations 112, 120, 121, 122, 123. Entities within the LAN 103 and the WAN 104 may be identified using network or protocol addresses such as, for example, ***Internet protocol (IP) addresses***, media access control (MAC) addresses, and electronic serial numbers (ESN's).”
- [0050] “Any of the PC's (105, 106, 107, 121) may indirectly access and control any of the plurality of media peripheral devices (108, 109). Such access and control may be accomplished in communication pathways via the MPS's (101, 102) or outside of the MPS's (101, 102). Similarly, any of the MPS's (101, 102) may indirectly access and control any of the media peripheral devices (108, 109) via a remote control 110 interacting with a TV screen 111, or outside thereof.”
- [0052]: “There are many types of indirect control commands available to manipulate the various media peripheral devices. Each media peripheral device may have particular indirect commands unique to that device or type of device. Many of the indirect commands will find parallels to the available direct commands. ***Exemplary commands include*** turning media peripheral devices 108 and/or 109 on and off, initiating play, capture,

select, erase/delete, zoom, focus, pan, tilt, fast forward, rewind, *scan*, list, skip, etc.”

- [0054] “FIG. 1B is a diagram illustrating an embodiment of a device view 130 with drop down menus for controlling the operation of media peripheral devices on the media exchange network 100 of FIG. 1A, in accordance with various aspects of the present invention. ***The device view 130 comprises a table of media peripheral device options 131 versus media content categories 132 (e.g., image files, video files, audio files, album titles)*** for each of the media peripheral devices. The media peripheral device options 131 listed in the device view 130 include a digital camera option 133, a digital camcorder option 134, a MP3 player option 135, a multi-media PDA option 136, a home juke-box option 137, a first home PC option 138, a mobile multi-media gateway option 139, and a second home PC option 140.”
- [0107]: “The MPS 900 comprises a media processing unit (MPU) 904, remote user interface(s) 905, and a TV screen 918 to provide integrated media processing capability and indirect user interface capability. The remote user interfaces 905 may comprise a ***voice or keyed remote control*** 906, keyboards and pads 907, a remote PC access interface 908, and a remote media system access interface 909 (i.e., providing access from another MPS).”

207. Karaoguz teaches a main device, such as a PC, and a “media peripheral device” such as an “MP3 player” that plays “audio files.” Karaoguz, [0050], [0054]. The PC and MP3 player are connected over a “communication network,” such as “the internet.” Karaoguz, [0024], [0047]. Karaoguz further discloses various “commands” for manipulating the media, including “scan.” Karaoguz, [0024], [0052]. Moreover, Karaoguz further discloses using “voice” to control operations. Karaoguz, [0107].

208. As described above in § IX.B.2, a person of ordinary skill in the art

would have been motivated to apply Karaoguz’s teachings of executing a “command” to scan “audio” to Victorian’s “ear-level device” (as modified by Jannard).

C. GROUND 3: Ground 1 in further view of Nathan (Claims 10-14):

1. Overview of Nathan

209. Nathan discloses a system to “efficiently perform various financial services for a client automatically, including complex financial transactions,” over “the internet,” using “speech signal spoken by a user.” [0006]-[0007], [0020]. Nathan’s system “includes a speech processor” that “perform[s] a requested financial service transaction.” *Id.*, [0007]. A “selection module 210 queries 702 the user to select an activity,” analyzes the user’s spoken response, and calls the appropriate module to carry out the financial transactions. *Id.*, [0035]. For example, if “the spoken term indicates...that the user confirmed the transaction the trade module 218 executes 1140 the order and the transaction is complete.” *Id.*, [0047]. *See also Id.*, [0021], [0022], [0026], [0027].

2. Motivation to Apply Nathan’s Teachings to Victorian and Jannard

Victorian, Jannard, and Nathan are analogous art. Victorian, Jannard, and Nathan are in the same field as the ’632 patent, which includes audio devices ’632, Abstract, 3:57-4:3; Victorian, [0016], [0049]; Jannard, [0276], [0292]; Nathan, [0004], [0007]; see §IX.A.3. Victorian, Jannard, and Nathan are reasonably pertinent

to alleged problems addressed by the '632 patent of using sound detection in an audio device to improve functionality. *E.g.*, '632 patent, Abstract (“The interaction element and logic circuit are configured to initiate control of audio content for output from the speaker in response to at least one voice command detected in captured audio.”); Victorian, [0030] (“In one embodiment, processor 460 includes a speech recognition module to enhance the audio signal received by and/or transmitted from ear-level device 110.”); Jannard, [0187] (“[T]he audio device 10C can wirelessly interact with an interactive audio device, such as a cellular phone, cordless phone, or a computer which responds to voice commands.”); Nathan, [0007] (“The system includes a speech processor that receives a speech signal spoken by a user into a telephone and identifies the term spoken by the user by comparing the received signal with a predefined vocabulary. The speech applications interpret the recognized term and, after receiving the necessary information, the speech applications perform a requested financial service transaction and provide information to the user regarding this transaction via the telephone system, for example.”).

A person of ordinary skill in the art would have been motivated to apply Nathan’s teachings of executing a “transaction” in response to a voice command, such as a user’s “spoken term,” to Victorian’s “ear-level device” (as modified by Jannard). For example, Jannard describes “internet browsing” capabilities,

responsive to “voice commands,” but does not enumerate all the possible internet browsing applications that can be implemented with voice commands. Jannard, [0292], [0299], [0301]. A person of ordinary skill in the art would have been motivated to look to analogous art—such as Nathan—to advantageously identify additional applications of internet browsing using voice commands, such as purchasing stocks. *See* Nathan, [0020]. Moreover, extending the internet-browsing capabilities to include financial transactions as taught by Nathan would be a simple and straightforward modification.

210. Thus, a person of ordinary skill in the art would have found it routine, straightforward and advantageous to apply Nathan’s teachings of executing stock transactions by voice command to Victorian’s ear-level device (as modified by Jannard’s teachings of a voice control interface) with internet browsing capabilities and would have known such a combination (yielding the claimed limitations) would predictably work and provide the expected functionality.

3. Claim [11]

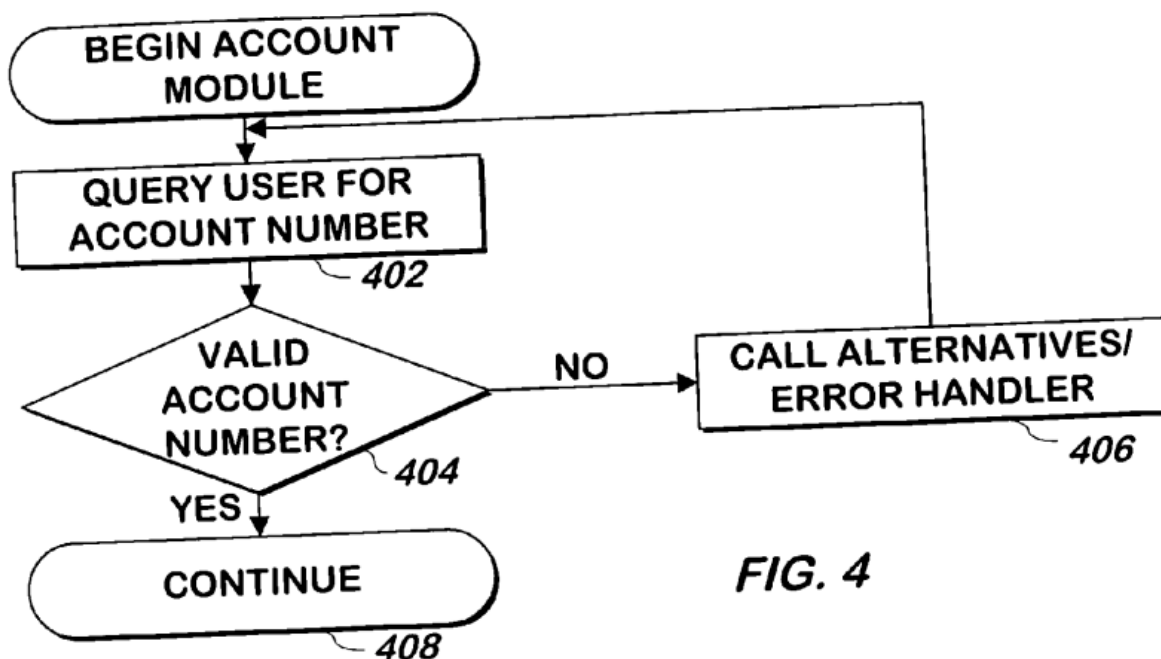
211. **Claim [11] recites “An earpiece according to claim 2, where the voice command is to link to a user's investment account.”** See my discussion of claim [2] above (§IX.A.4.m-n). The limitation added by the claim is met by Victorian in view of Jannard and Nathan.

212. As I explain above in connection with elements [1.h.1]-[1.h.5] and

[4.h] Jannard discloses voice commands provided via user controls 322 (§IX.A.4.i-1, § IX.A.4.t).

213. Nathan discloses the voice command (e.g., “the term spoken by the user” “corresponds to a valid account number”) is to link a user’s investment account (e.g., “[i]f the spoken term is a valid account number the process continues 408” and allows the user to access their account). See, e.g., Nathan:

- Fig. 4



- [0031]: “FIG. 4 is a flow chart illustrating a process performed by the account module 204 according to a preferred embodiment of the present invention. *The account module 204 queries 402 the user for the user’s account number using the procedure described above, for example.* The user speaks into the telephone 101 and the VoiceBroker™ system 100 identifies the term spoken by the user. *The account module 204 then determines 404 if the spoken term corresponds to a valid account number by accessing the customer database 142.* If the spoken term is a

valid account number the process continues 408 and process control returns to the controller 202. If the spoken term is not a valid account number the account module 204 calls 406 the alternatives/error handler (A/EH) 206.”

- [0032]: “FIG. 5 is a flow chart illustrating a process performed by the alternatives/ error handler 206 according to a preferred embodiment of the present invention. The alternatives/error handler 206 manages the procedure for handling spoken terms that do not answer the most recent query. For example, in the preferred embodiment the user has the option of requesting help in answering a query. In response to a “help” request the alternatives/error handler 206 will generate a message that will provide more information to the user that is related to the question asked. This is one type of alternate option that is available to the user. In some contexts other alternate options are also available to the users. For example, after the user has satisfied the security module 208 the user can return to a main menu, e.g., the selection module 210, by speaking the words “main menu” at any time. The user can similarly transfer to a specific module to perform a specific function within the VoiceBroker™ system 100 by speaking a predetermined term, e.g., “trade.” The user may also modify the user's security code at any time after successfully completing the security module 208. Another option enables the user to transfer to a customer service representative by saying the phrase “customer service” or a similar term/phrase, e.g., “representative”, “agent”, “service”, “transfer”, or “broker”. If the alternatives/error handler 206 determines 502 that the spoken term is a valid alternate option then the alternatives/error handler 206 calls the appropriate function to implement 504 the function. If the spoken term is not a valid alternate option the alternatives/error handler 206 informs 506 the user that the spoken term is an invalid response and determines 508 if the user should be transferred to a customer service representative. The user is automatically transferred to a customer service representative if the user does not answer a question correctly after a number of attempts. The number of attempts can vary depending upon the context of the system, e.g., if the process of the security module 208 is being implemented then the number of incorrect responses that will cause the alternatives/error handler 206 to transfer the user to customer service may be less than other, less critical, contexts. If the alternatives/error handler 206 determines 508 that the user should be transferred to a customer service representative, the alternatives/error handler 206 transfers 510 the user. If

the alternatives/error handler 206 determines 508 that such a transfer is not necessary than the process continues and control passes to the module that called the alternatives/error handler 206. For example, if the account module 204 calls 406 the alternatives/error handler 206, when the alternatives/error handler 206 continues 512, the account module 204 will again query 402 the user for the account number and the process continues until either a correct account number is spoken or the alternatives/error handler 206 transfers 510 the user to a customer service representative. *As described above, if the spoken term identifies a valid account number the account module 204 continues 408 and the controller 202 calls 304 the security module 208. The operation of the security module 208 is now described.*”

214. A person of ordinary skill in the art would have understood that, by allowing the user to continue to access their account, the user has been linked to their account, and that the account number is a command to make that link. This is consistent with the disclosure of the ’632 patent, which never mentions this concept outside of claim 11. In the context of claim 11, there is no additional entity besides the “user” or the user’s device that would be “linked” to the investment account. In this regard, the phone (or whatever device the user is using) is also linked to the account, as the user could pass the phone (or other device) to another person and the account would still be linked. In the context of the combination with Victorian, Victorian’s earpiece would be linked to the account.

215. As described above in §IX.C.2, a person of ordinary skill in the art would have been motivated to apply Nathan’s teachings of connecting a user to an account to Victorian’s “ear-level device” (as modified by Jannard).

4. Claims [10] and [12]

216. Claim [10] recites “An earpiece according to claim 2, where the voice command is to search for a particular stock value” and claim [12] recites “An earpiece according to claim 11, where a second voice command is to search for a particular stock value.” See my discussions of claim [2] and claim [11] above (§IX.A.4.m-n, § IX.C.3). The limitation added by the claim is met by Victorian in view of Jannard and Nathan.

217. As I explain above in connection with elements [1.h.1]-[1.h.5] and [4.h] Jannard discloses voice commands provided via user controls 322 (§IX.A.4.i-1, § IX.A.4.t).

218. Nathan discloses the voice command (e.g., “a speech signal spoken by a user”) is to search for a particular stock value (e.g., “providing stock quotes”).

See, e.g., Nathan:

- [0007]: “The system includes *a speech processor that receives a speech signal spoken by a user into a telephone and identifies the term spoken by the user by comparing the received signal with a predefined vocabulary. The speech applications interpret the recognized term and, after receiving the necessary information, the speech applications perform a requested financial service transaction and provide information to the user regarding this transaction via the telephone system, for example. Some of the financial service transactions supported by the present invention include, providing stock quotes, providing mutual fund quotes, providing option quotes, providing account information, trading stocks, trading mutual funds, and trading options. In addition, the present invention utilizes the speech recognition technology to provide a security system that prevents unauthorized access to a financial account.*”

219. Nathan teaches “speech applications [that] perform a requested financial service transaction” such as “providing stock quotes.” Nathan, [0007]. Nathan’s system teaches using speech to search for a particular stock value.

220. As described above in § IX.C.2, a person of ordinary skill in the art would have been motivated to apply Nathan’s teachings of providing stock quotes to Victorian’s “ear-level device” (as modified by Jannard).

5. Claims [13] and [14]

221. **Claim [13] recites “An earpiece according to claim 12, where a third voice command is to buy the particular stock” and claim [14] recites “An earpiece according to claim 12, where a third voice command is to sell the particular stock.”** See my discussion of claim [2] above (§IX.A.4.m-n). The limitation added by this claim is met by Victorian in view of Jannard and Nathan.

222. As I explain above in connection with elements [1.h.1]-[1.h.5] and [4.h] Jannard discloses voice commands provided via user controls 322 (§IX.A.4.i-l, § IX.A.4.t).

223. Nathan discloses a third voice command (*e.g.*, “spoken term”) is to buy a stock (claim 13) or sell a stock (claim 14) (*e.g.*, “spoken term indicates that the user wants to perform a trade” including “buy or sell”). *See, e.g.*, Nathan:

- [0043]: “As described above, if, when the selection module 210 has process control, the spoken term indicates that the user wants to perform a trade, the selection module 210 calls 714 the trade module 218. The process performed by the trade module 218 is now described. FIGS.

11A-D are flow charts illustrating a process performed by a trade module 218 according to a preferred embodiment of the present invention. If the user requests to perform a trade the trade module 218 queries 1102 the user on the name of the financial instrument to trade, e.g., stock, mutual fund, option, future, commodity, unit investment trust (UIT), real estate investment trust (REIT), foreign currency, real property, or precious metal, for example. The trade module 218 determines 1104 if the spoken term represents a valid financial instrument. If not, the trade module 218 calls 1106 the alternatives/error handler 206. If the spoken term does represent a valid financial instrument the trade module 218 obtains 1108 quote information for the financial instrument. Some techniques for obtaining quote information are described above. *The trade module 218 then informs 1110 the user of the quote information and asks 1112 the user for the trade type, e.g., buy or sell. If the received spoken term indicates 1114 that the user wants a quote on another financial instrument the trade module 218 returns to step 1102. If not, the trade module 218 determines whether the spoken term indicates that the user wants to either buy 1116 or sell 1118. If the spoken term does not correspond to either buying or selling the trade module 218 calls 1120 the alternatives/error handler.*

224. Nathan discloses that a user executes a “buy or sell” trade for a “financial instrument” such as a “stock” by using a “spoken term.” Nathan, [0043].

225. As described above in § IX.C.2, a person of ordinary skill in the art would have been motivated to apply Nathan’s teachings of buying and selling stocks to Victorian’s “ear-level device” (as modified by Jannard).

D. GROUND 4: Ground 1 in further view of Camp (Claims 16-17):

1. Overview of Camp

226. Camp is directed to the measurement and control of sound pressure levels in an ear canal. “The present invention automatically controls the SPL in the ear canal 14 by measuring a current SPL in the ear canal 14 and adjusting the

volume of projected audible signals based on the measured SPL.” Camp, [0014].

Camp discloses that such measurements are important because “elevated pressure deviations may damage the ear.” Camp, [0013].

2. Motivation to Apply the Camp’s Teachings to Victorian and Jannard

Victorian, Jannard, and Camp are analogous art. Victorian, Jannard, and Camp are in the same field as the ’632 patent, which includes personal audio devices ’632, Abstract, 3:57-4:3; Victorian, [0016], [0049]; Jannard, [0276], [0292]; Camp, [0002], [0011]; see §IX.A.3. Victorian, Jannard, and Camp are reasonably pertinent to alleged problems addressed by the ’632 patent of using sound detection in an audio device to improve functionality. ’632, Abstract (“The interaction element and logic circuit are configured to initiate control of audio content for output from the speaker in response to at least one voice command detected in captured audio.”); Victorian, [0030] (“In one embodiment, processor 460 includes a speech recognition module to enhance the audio signal received by and/or transmitted from ear-level device 110.”); Jannard, [0187] (“[T]he audio device 10C can wirelessly interact with an interactive audio device, such as a cellular phone, cordless phone, or a computer which responds to voice commands.”); Camp, [0014] (“The present invention automatically controls the SPL in the ear canal 14 by measuring a current SPL in the ear canal 14 and adjusting the volume of projected audible signals based on the measured SPL.”), Abstract, [0004], [0031].

A person of ordinary skill in the art would have been motivated to apply Camp's teaching of detecting the sound pressure level (including sound pressure over a period of time) to Victorian's ear-piece (as modified by Jannard). A conventional earbud "channels most of the audible signals projected by earbud 20 directly to ear drum 16. This feature typically provides superior sound quality relative to other conventional headphones. However, this feature also produces higher pressure deviations, referred to herein as sound pressure levels (SPLs), in the ear canal 14 when compared to other non-earbud headphones operating at the same volume. These elevated pressure deviations may damage the ear." Camp, [0013]. This functionality is implemented in Camp's processor (DSP). Camp, [0018] ("FIG. 4 illustrates another exemplary threshold detection process 205 that may be implemented by DSP 124.") A person of ordinary skill in the art thus would have been motivated to apply Camp's teachings of "automatically adjusting the volume of a headset" to Victorian's ear-piece (as modified by Jannard) in order to advantageously prevent damage to a user's ear. Camp, Abstract. A person of ordinary skill in the art further would have understood that applying these teachings to Victorian would have involved simply reprogramming Victorian's processor to keep track of SPL.

Thus, a person of ordinary skill in the art would have found it routine, straightforward and advantageous to apply Camp's teachings of calculating the

sound pressure level dosage of a user's ear to Victorian's ear-level device (as modified by Jannard's teachings of a voice control interface) and would have known such a combination (yielding the claimed limitations) would predictably work and provide the expected functionality.

3. Claim [16]

227. Claim [16] recites “An earpiece according to claim 15, further including the operations of: measuring the sound pressure level using the ear canal microphone.” See my discussion of claim [15], above (§IX.A.4.y). As I explain above in Section IX.A.4.z, Victorian teaches the limitation added by this claim. To the extent further disclosure is required, the limitation added by this claim is met by Victorian in view of Camp.

228. Camp discloses measuring the sound pressure level using an ear canal microphone (*e.g.*, “pressure transducer 114...configured to detect sound pressure deviations”). *See, e.g.*, Camp:

- [0015]: “Each earbud 110 includes speaker 112 and pressure transducer 114. Speaker 112 may comprise any speaker conventionally used in earbud headsets, while transducer 114 may comprise any transducer configured to accurately detect sound pressure deviations. When earbud 110 is disposed in an ear canal 14, speaker 112 projects audible signals into ear canal 14, causing pressure deviations in the ear canal 14. ***Transducer 114 senses these pressure deviations, and converts the sensed pressure deviations to an electrical signal representative of the SPL in the ear canal 14.*** As used herein, SPL refers to an analog or digital electrical signal used in an electronic system or computer program that is representative of the physical SPL present in ear canal 14. The measured SPL may be the result of the projected audible signal from

speaker 112, external environmental noise coupled to ear canal 14, or any combination thereof. According to one exemplary embodiment, transducer 114 and speaker 112 are acoustically coupled to each other in the outer ear canal 14 and acoustically isolated from each other in earbud 110 to ensure that the measured SPL corresponds to the SPL in the ear canal 14.”

229. Camp discloses “earbuds 110” with “transducer 114 and speaker 112 are acoustically coupled to each other in the outer ear canal 114,” and transducer 114 measures a sound pressure level (SPL).

230. As described above in §IX.D.2, a person of ordinary skill in the art would have been motivated to apply Camp’s teachings of measuring SPL with a microphone in the ear canal to Victorian’s “ear-level device” (as modified by Jannard).

4. Claim [17]

231. **Claim [17] recites “An earpiece according to claim 15, further including the operations of: calculating the sound pressure level dosage of a user using the measured sound pressure level.”** See my discussion of claim [15], above (§IX.A.4.y). The limitation added by this claim is met by Victorian in view of Camp.

232. Victorian discloses the earpiece (*see* [1.pre]).

233. Camp discloses calculating the sound pressure level dosage of a user using the measured sound pressure level (*e.g.*, “the detected SPL value exceeds the threshold (block 220) for more than a predetermined length of time”). *See, e.g.*,

Camp:

- Fig. 4

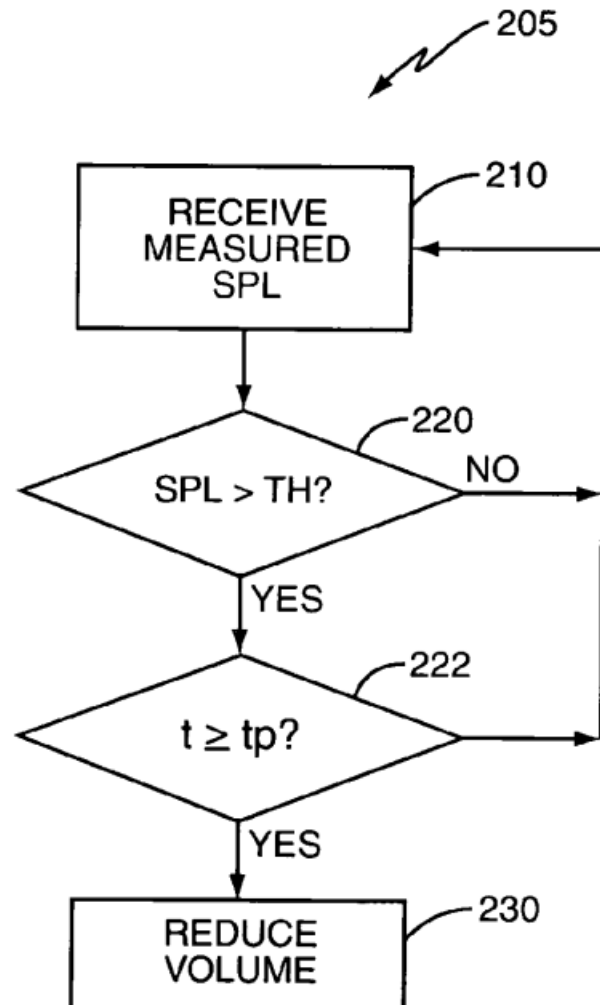


FIG. 4

- [0018]: “FIG. 4 illustrates another exemplary threshold detection process 205 that may be implemented by DSP 124. *After receiving a measured SPL (block 210), DSP 124 detects a peak or RMS value of the measured SPL and compares the detected SPL value to a predetermined threshold (block 220). If the detected SPL value exceeds the threshold (block 220) for more than a predetermined length of time (block 222), DSP 124 generates a control signal 136 (block 230) to*

reduce the volume. For example, if the detected SPL exceeds 100 dBA for more than 60 minutes or exceeds 65 dBA for more than 40 hours in one week, control signal 136 directs controller 130 to reduce the volume, and therefore, to reduce the SPL in the ear canal 14 to an acceptable level. Otherwise, DSP 124 continues to monitor the SPL relative to the predetermined threshold and time limit (blocks 210, 220, 222). It will be appreciated that the present invention is not limited to the single threshold and time limit of the above examples. In alternative embodiments, DSP 124 may track multiple time intervals relative to multiple different SPL thresholds. For example, a first timer may track how long the detected SPL exceeds a first threshold, such as 75 dB, while second and third timers may track how long the detected SPL exceeds second and third thresholds, respectively. Based on these thresholds and the pre-determined time limits associated with each timer, DSP generates a volume control signal 136 that controls the volume of the projected audible signals.”

234. Camp discloses “earbuds 110” calculate the SPL produced by “speaker 112” over a period of time, for example, “if the detected SPL exceeds 100 dBA for more than 60 minutes or exceeds 65 dBA for more than 40 hours in one week.” Camp, [0014], [0018]. If these amounts of SPL are exceeded, the device “reduce[s] the volume” to an acceptable level so as not to damage the user’s ear. Camp, [0018], [0013]. The extended measurement of SPL provides an SPL dosage of the user.

235. As described above in §IX.D.2, a person of ordinary skill in the art would have been motivated to apply Camp’s teachings of determining cumulative SPL in the ear canal to Victorian’s “ear-level device” (as modified by Jannard).

X. SECONDARY CONSIDERATIONS

236. I have seen no evidence in the '632 patent's prosecution history or elsewhere supporting any secondary considerations arguments, or evidence of nexus to any of the Challenged Claims. See generally '632 File History. Indeed, per my discussion of the prior art referenced herein, any purported solutions to problems or unexpected results in the '632 patent were already well known. To the extent ST CasesTech, LLC asserts the existence of any secondary considerations in its responses, I reserve my right to address any such evidence.

XI. CONCLUSION

237. In summary, I have concluded that the Challenged Claims are rendered obvious over the prior art references as discussed above. Specifically, in my opinion, for at least the reasons above, the Challenged Claims of the '632 patent obvious.

238. To the extent it is argued that any further disclosure is required for an element in the Challenged Claims that I have identified as being disclosed, a person of ordinary skill in the art would certainly have found that element obvious to include based on the sample explicit disclosure or inherent disclosures and analysis I have identified above.

239. I reserve the right to supplement my opinions in the future to respond to any arguments that Patent Owner or its expert(s) may raise and to take into

account new information as it becomes available to me.

240. In signing this declaration, I recognize that the declaration will be filed as evidence in a contested case before the Patent Trial and Appeal Board of the United States Patent and Trademark Office. I also recognize that I may be subject to cross-examination in the case and that cross-examination will take place within the United States. If cross-examination is required of me, I will appear for cross-examination within the United States during the time allotted for cross-examination.

241. I reserve the right to supplement my opinions in the future to respond to any arguments that Patent Owner or its expert(s) may raise and to take into account new information as it becomes available to me.

242. I hereby declare that all statements made herein of 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 of Title 18 of the United States Code.

243. Executed this 1st day of September 2024, in Lexington KY.



Dr. Michael T. Johnson

APPENDIX A

APPENDIX B

Exhibit ("Ex.")	Description
1001	U.S. Patent No. 11,521,632 ("632 patent")
1002	File History of U.S. Application No. 16/785,822 ("632 File History")
1005	U.S. Provisional Application No. 60/806,769
1006	European Pub. No. EP1855400A1 ("Srail")
1007	European Pub. No. EP1519625A2 ("Victorian")
1008	U.S. App. Pub. No. 2006/0132382 ("Jannard")
1009	U.S. App. Pub. No. 2004/0117846 ("Karaoguz")
1010	U.S. App. Pub. No. 2004/0167847 ("Nathan")
1011	U.S. App. Pub. No. 2007/0274531 ("Camp")
1013	U.S. App. Pub. No. 2003/0138118A1 ("Stahl")
1014	U.S. Patent Application Publication No. 2006/0136378 ("Martin")
1015	Spoken language Processing, Huang, 2001 ("SLP")
1017	Michael H. Cohen, Voice User Interface Design (2004) ("Cohen")
1018	U.S. App. Pub. No. 2006/0116175A1 ("Chu")
1019	U.S. App. Pub. No. 2002/0068558A1 ("Janik")
1020	U.S. App. Pub. No. 2005/0240236A1 ("Daum")
1021	U.S. App. Pub. No. 2005/0125152A1 ("Fuchs")
1022	Computer Dictionary, 2 nd Edition, 1994, excerpt page 365.
1023	U.S. App. Pub. No. 2006/0262938A1 ("Gauger")

Exhibit ("Ex.")	Description
1024	JAE SOO LIM, ENHANCED BANDWIDTH COMPRESSION OF NOISY SPEECH BY ESTIMATION OF SPEECH AND ITS MODEL PARAMETERS (1978) ("Lim")