

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

HARMAN INTERNATIONAL INDUSTRIES, INCORPORATED,

Petitioner,

v.

ST CASESTECH LLC,

Patent Owner.

Case IPR2024-01301

U.S. Patent No. 8,805,692

**DECLARATION OF DR. MICHAEL T. JOHNSON IN SUPPORT
OF PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 8,805,692**

TABLE OF CONTENTS

	Page
I. INTRODUCTION	1
II. QUALIFICATIONS	4
III. LEGAL UNDERSTANDING	8
IV. BACKGROUND OF THE TECHNOLOGY	14
V. THE '692 PATENT	15
A. '692 Patent Overview	15
B. The Prosecution History	18
VI. LEVEL OF ORDINARY SKILL IN THE ART	21
VII. CLAIM CONSTRUCTION	22
VIII. GROUNDS OF UNPATENTABILITY	23
IX. SPECIFIC GROUNDS FOR CHALLENGE	24
A. GROUND 1 (Jannard in view of Wittenberg, Claims 1-11, 13,14, 16-20).....	25
1. Overview of Jannard	25
2. Overview of Wittenberg	28
3. Motivation to Apply Wittenberg's Teachings to Jannard.....	30
4. Invalidity of Claims 1-11, 13, 14, 16-20 Over Jannard in view of Wittenberg.....	33
B. GROUND 2: GROUND 1 IN FURTHER VIEW OF RELAN (Claim 12):.....	117
1. Overview of Relan	117
2. Motivation to Apply the Relan's Teachings to Jannard and Wittenberg.....	118
3. Claim 12.....	121

C.	GROUND 3: GROUND 1 IN FURTHER VIEW OF BARNES (Claim 15):	124
1.	Overview of Barnes	124
2.	Motivation to Apply Barnes’ Teachings to Jannard and Wittenberg.....	124
3.	Claim 15.....	126
X.	SECONDARY CONSIDERATIONS	128
XI.	CONCLUSION	129
	APPENDIX A.....	131
	APPENDIX B	132

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. 8,805,692 (“the ’692 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 1-20 of the ’692 patent (“Challenged Claims”) are obvious over the prior art references identified below.

2. I understand that the ’692 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 ’692 patent. I understand that the ’692 patent has been provided as Ex. 1001.

3. I understand that the file history of the ’692 patent has been provided as Ex. 1002. I have considered this file history, and I will refer to it as “the ’692 File History.”

4. I understand that the ’692 patent is a continuation of 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 (1017). I have considered this file history.

5. I understand that the '692 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 '692 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 '692 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 '692 patent; my own knowledge and

experiences gained from my work experience in the field of the '692 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.

- **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 1004.
- **U.S. Patent No. 7,529,677 B1** to John W. Wittenberg (“Wittenberg”), titled “Methods and apparatus for remotely processing locally generated commands to control a local device” filed on Jan. 21, 2005, and issued May 5, 2009. I understand that a copy of this reference has been designated as Exhibit 1005.
- **U.S. Patent Application Publication No. 2006/0020960** to Sandeep Relan, Brajabandhu Mishra, and Rajendra Khare (“Relan”), titled “System, method, and apparatus for secure sharing of multimedia content across several electronic devices” filed on Sep. 6, 2005, and published Jan. 26, 2006. I understand that a copy of this reference has been designated as Exhibit 1007.
- **U.S. Patent Application Publication No. 2003/0220835** to Melvin Barnes (“Barnes”), titled “System, method, and computer program product for providing location based services and mobile e-commerce” filed on May 23, 2002, and published Nov. 27, 2003. I understand that a copy of this reference has been designated as Exhibit 1008.

9. I understand that there is a pending district court litigation between Petitioner and ST CasesTech, LLC (“Harman Litigation”) involving the '692

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. 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). 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.

23. 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

24. 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.

25. 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

26. 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.

27. 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.

28. 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.

29. 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.

30. 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

31. As discussed in Section V, the '692 patent is directed to a wearable device with an ambient microphone in or on its housing that captures voice commands and other ambient audio. E.g., '692 patent, Abstract, 9:13-21. The voice controlled interface uses voice commands to control the audio of the device. '692 patent, 11:33-54. For example, in one embodiment, a voice command results in the skipping to the next piece of audio content stored in the memory of the wearable device. '692 patent, 11:33-54. The purported inventive concept of the '692 patent is that the wearable device controls media of the wearable device in response to voice commands captured at the ambient microphone and analyzed at a server. '692 patent, Abstract.

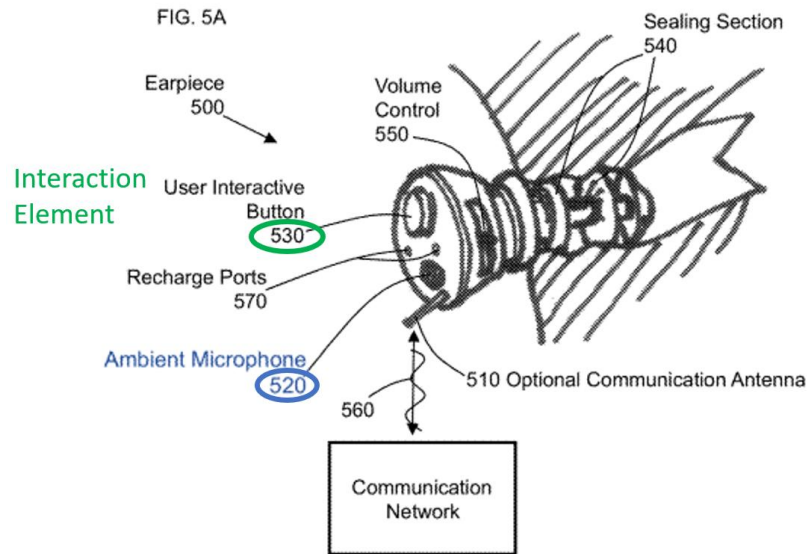
32. The technology at issue in the '692 patent was developed by others before the '692 patent. For example, Jannard teaches an audio device that is wearable, and receives voice commands that result in the “fast-forward, skip, cue, play, [and] pause” of music stored on the device. *See e.g.*, Jannard, [0257], [0265]. Further, Wittenberg teaches a remote server used to carry out the speech recognition of voice commands entered at the local device and transmits appropriate instructions back to a local device. Wittenberg, 2:35-64.

V. THE '692 PATENT

A. '692 Patent Overview

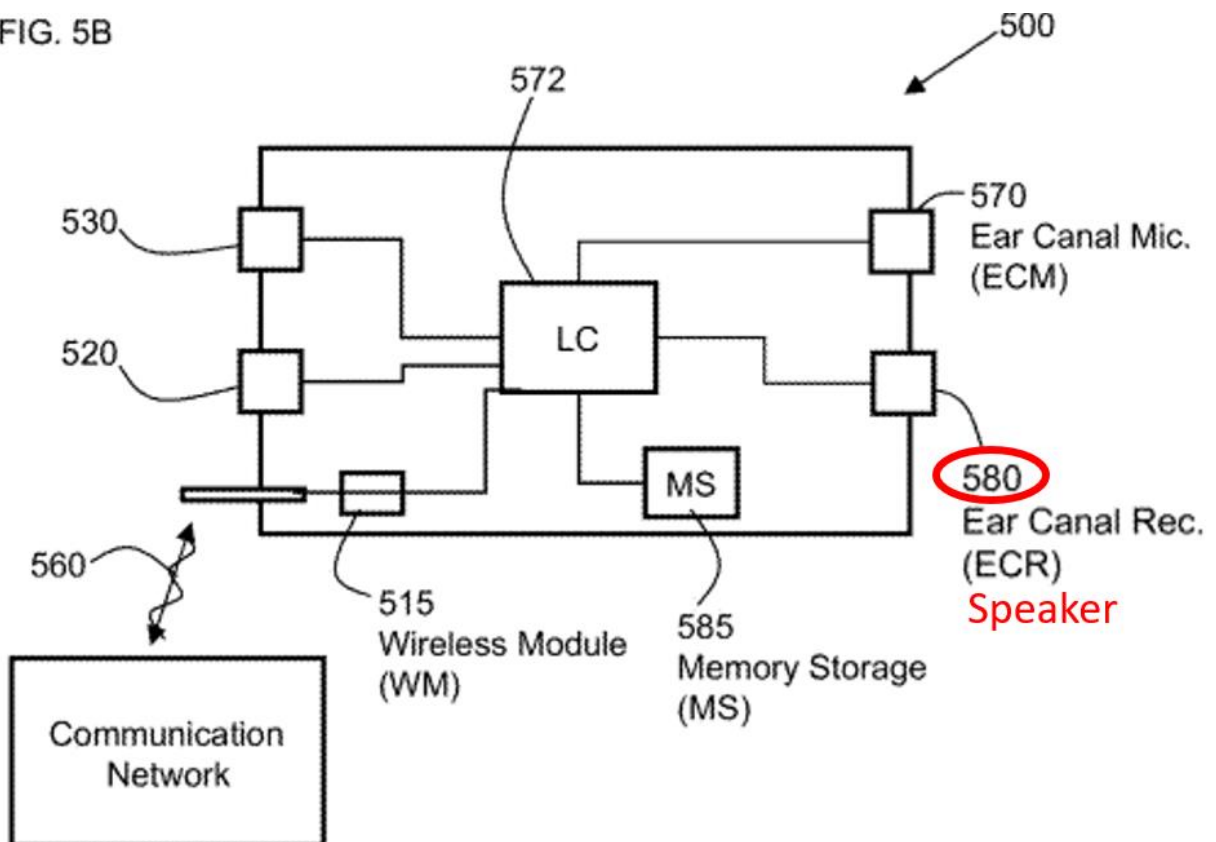
33. The '692 patent describes “methods and devices for the storage and recall of audio content via an earpiece.” '692 patent, 1:14-18. More particularly, the '692 patent describes methods “for the storage and playing of music or verbal content on a system that is built into a headphone.” '692 patent, 1:14-18. The '692 patent attempts to solve the issue of compatibility combining audio content players and earpieces systems; “[c]ombining the capacity for audio download and playing in an earpiece system” with a “user interaction system.” '692 patent, 1:30-34. Further, '692 describes the “download of audio content into an earpiece.” '692 patent, 1:35-36.

34. The '692 patent describes a wearable device, with an embodiment as an earpiece. '692 patent, 2:15-27. The '692 patent includes as an example of “earpiece” as “e.g., earbuds [and] headphones.” '692 patent, 2:15-27. The wearable device has several elements that are operatively connected to a logic circuit, including: an **ambient microphone**, an **ear canal receiver** (or “**speaker**”), a communication module, a memory storage unit, and a user **interaction element**. '692 patent, 8:33-37, 9:22-25, 39-43, Figs. 5A, 5B. The memory storage unit “can store data such as registration, preference, and audio content data.” '692 patent, 9:21-27.



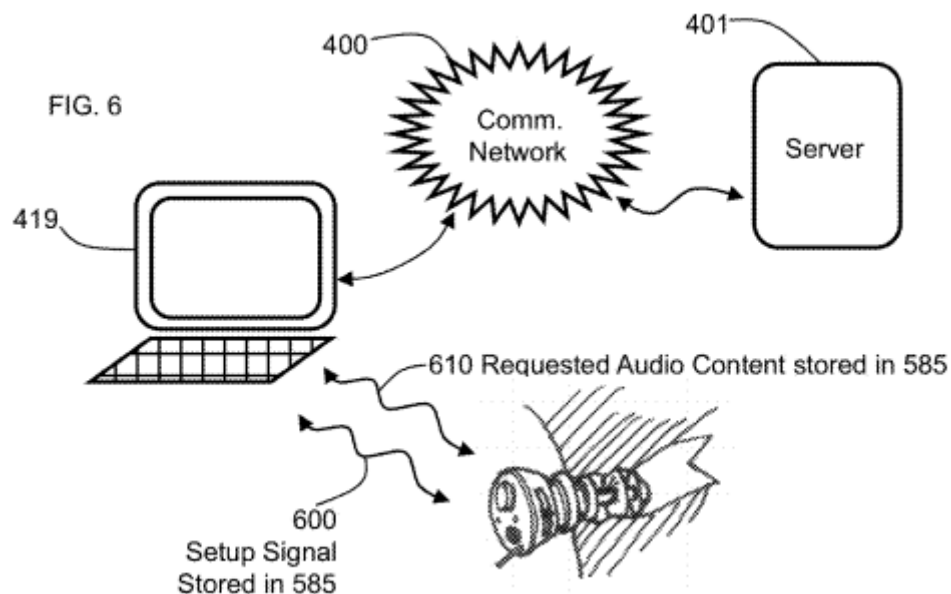
'692 patent, Fig. 5A

FIG. 5B



'692 patent, Fig. 5B

35. A communication module supports “bidirectional” transmission protocols (*e.g.*, “TCP/IP, USB, IEEE 1394, IEEE 802.11, Bluetooth, A2DP, GSM, CDMA or others”) via a communication network that allows “communication between the earpiece 500 and a communication network.” ’692 patent, 5:32-36, 8:12-15, 9:9-12. The user “use[s] his/her computer 419 to communicate with a server 401 (*e.g.*, via a communication network 400).” ’692 patent, 9:28-33. The “information can then be transmitted to set the stored parameters in the memory storage device 585 of the earpiece 500.” ’692 patent, 9:29-38, Fig. 6. Further, “audio content can be downloaded 610 into the memory storage device 585 of the earpiece 500.” ’692 patent, 9:29-38, Fig. 6.



’692 patent, Fig. 6

36. An **ambient microphone (AM)** is “used to pick up ambient audio

content,” such as voice commands. ’692 patent, 9:13-21, 11:32-49. The audio content can be recorded “to the data memory storage system.” ’692 patent, 12:55-61; *see also* 9:13-21.

37. The user may initiate communication with the server system via “voice commands.” ’692 patent, 4:39-41. The ’692 patent offers an example of a user searching the web using voice commands and receiving results via a text to speech synthesizer. ’692 patent, 4:39-41. The ’692 patent describes a “user **interaction element 530** (e.g., a **button**) that can be used to turn the earpiece 500 on, or if on then activate an audio play command to start playing saved audio content.” ’692 patent, 8:67-9:3.

B. The Prosecution History

38. The ’692 patent application (No. 14/148,752) was filed January 7, 2014. The ’692 patent was issued on August 12, 2014. The ’692 patent claims priority from application No. 11/774,965, filed on July 9, 2007, which claims priority to Provisional application No. 60/806,769, filed on July 8, 2006.

39. During prosecution, the examiner issued a rejection for Claims 1, 2, 4, 6-11, 14, and 18-20 under pre-AIA 35 U.S.C. 102(b) as being anticipated by U.S. Patent Application Publication No. 2003/0130852 (“Tanaka”) (Ex. 1009). ’692FH, 144-46. The examiner issued a rejection for Claims 5, 12-13, and 15-17 under pre-AIA 35 U.S.C. 103(a) as being obvious over Tanaka in view of U.S. Patent

Application Publication No. 2010/0311390 (“Black”) (Ex. 1010). ’692FH, 147-49. The examiner issued a rejection of Claims 3 and 12 under pre-AIA 35 U.S.C. 103(a) as being unpatentable over Tanaka in view of U.S. Patent Application Publication No. 2007/0165875 (“Rezvani”) (Ex. 1011). ’692FH, 149.

40. For Claim 1, Applicant argued that “Tanaka fails to disclose or suggest a wearable device and an ambient microphone within or on the housing for the wearable device and a logic circuit communicatively coupled to the ambient microphone.” ’692FH, 169. Applicant argued Tanaka “disclose[s] and illustrate[s] a boom microphone that is a directional and not an ambient microphone.” ’692FH, 169. Applicant further argued that Tanaka “in numerous instances teaches away from using an ambient microphone since Tanaka readily concedes that noises would be ‘superposed onto the speech signal detected by the microphone and the recognition performance is lowered.’” ’692FH, 171, *see also* 170-173, 175. Regarding Claim 18, Applicant argued that Tanaka fails to disclose an ambient microphone, and thus is allowable. ’692FH, 175. Applicant made similar arguments for claim 19 and 20. ’692FH, 175.

41. For claim 2, Applicant stated that Tanaka fails to disclose an interaction element and logic circuit configured to initiate control of access to the media content played back at the wearable device. ’692FH, 173. Applicant argued that Tanaka fails to disclose that its speech recognition device controls access to

media content at the wearable device. '692FH, 173-175. Applicant argued that claims 3 and 4 are allowable as being dependent on an allowable base claim 1, and claim 2. '692FH, 174-175.

42. Applicant made arguments to support the remaining dependent claims. '692FH, 173-79.

43. A final rejection was issued by the examiner for Claims 1, 2-4, 6-12, 14, and 18-20 under pre-AIA 35 U.S.C. 103(a) as being obvious by Tanaka in view of Rezvani. '692FH, 186. The Examiner stated that while "Tanaka doesn't teach where the microphone is ambient microphone," Rezvani teaches "a multimedia wireless headset comprising speech recognition system with a microphone for operating the wireless headset, where the microphone comprises ambient microphone (Fig.4; Par.0026)." '692FH, 186. The Examiner stated that the "use of ambient microphone in Tanaka's teaching would be obvious, to one of ordinary skill in the art, for the purpose of cancelling background noise as disclosed by Rezvani. '692FH, 187. The examiner stated that the use of an ambient microphone is "a newly added feature forcing the examiner to make changes." '692FH, 192. Further, the examiner "disagrees with applicant's assertion that Tanaka teaches away from using ambient microphone." '692FH, 192. "Tanaka simply chooses to use different kind of microphone that could simply be replaced with the well known ambient microphone as shown by Rezvani." '692FH, 192.

44. In response, Applicant amended each independent claim to recite: “wherein the logic circuit is configured to cooperatively initiate control of media content of the wearable device in response to analysis of audio captured at the wearable device at the ambient microphone and analyzed at a remote server in response to instructions received from the remote server” (claim 1) or “wherein the one or more processors is configured to cooperatively initiate control of media content of the wearable device in response to analysis of audio captured at the wearable device at the ambient microphone and analyzed at a remote server in response to instructions received from the remote server” (claims 18 and 20). ’692FH, 196, 200-201.

45. Claims 1-20 were then allowed. ’692FH, 233.

VI. LEVEL OF ORDINARY SKILL IN THE ART

46. 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.

47. 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.

48. Based on my knowledge and experience in the field and my review of the '692 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.

49. 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

50. I have been instructed that the Challenged Claims of the '692 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.

51. I understand that the parties have not proposed constructions for the '692 patent in the Harman Litigation.

VIII. GROUNDS OF UNPATENTABILITY

52. As I explain above in Section IV and V, a wearable device with an ambient microphone that captures and responds to voice commands was well-known in the art.

53. 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
Jannard	1004	U.S. 2006/0132382	6/22/2006
Wittenberg	1005	U.S. 7,529,677	5/5/2009
Relan	1007	U.S. 2006/0020960	1/26/2006
Barnes	1008	U.S. 2003/0220835	11/27/2003

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

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

- **Ground 1:** Challenged Claims 1-11, 13-14, and 16-20 of the '692 patent are rendered obvious by Jannard in view of Wittenberg.
- **Ground 2:** Challenged Claim 12 of the '692 patent is rendered obvious by Jannard in view of Wittenberg and Relan.
- **Ground 3:** Challenged Claim 15 of the '692 patent is rendered obvious by Jannard in view of Wittenberg and Barnes.

55. These Grounds above are not substantively the same as the art or arguments raised in the prosecution of the '692 patent. None of the references of the above grounds was considered during prosecution. And each of these references teaches limitations that the Examiner erroneously believed were missing from the prior art. For example, unlike the art cited during prosecution, including Tanaka and Rezvani, Jannard teaches a system with voice-activated electronic controls of media of a wearable device (Jannard, [0177], [0315]) and Wittenberg discloses offloading speech processing and resulting command formation to a remote device in contact with a local device (Wittenberg, 2:35-65). Relan discloses a wearable device capable of forming a peer-to-peer network, and Barnes discloses using voice commands for carrying out commercial transactions. Relan, [0059]; Barnes, [0124], [0128]. Indeed, the Examiner did not recognize any art teaching voice-command control of media content of an audio device.

IX. SPECIFIC GROUNDS FOR CHALLENGE

56. As I explain above in Section IV and V, although the '692 patent purports to have invented a method of storage and recall of audio content through an earpiece, such methods were well-known in the art. As described below, the cited prior art provides these teachings.

57. In this section, I explain why claims 1-11, 13-14, 16-20 of the '692 patent are rendered obvious by the teachings of Jannard (Ex. 1004) in view of

Wittenberg (Ex. 1005), why claim 12 is rendered obvious by the teachings of Jannard in view of Wittenberg and Relan (Ex. 1007), and why claim 15 is further rendered obvious by the teachings of Jannard in view of Wittenberg and Barnes (Ex. 1007). 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 (Jannard in view of Wittenberg, Claims 1-11, 13,14, 16-20)

1. Overview of Jannard

58. Jannard discloses a “wearable electronically enabled” device. Jannard, [0002]. Like the ’692 describes the issues of “audio content player...separate[] from the earpiece” (’692 patent, 1:26-30), Jannard recognizes the issue of an audio playback device that is separate from an audio interface. Jannard [0004]. Jannard describes the problem of the user having to “adjust or remove the audio interface” from the user’s ears whenever the user “wants to receive a cell phone call or even wear glasses or sunglasses.” Jannard, [0004]. Jannard addresses this problem by providing a wearable “audio device” with “a data port 308,” “**processor 312**,” “speakers 318,” “**user controls 322**,” and “**memory module 330**.” Jannard, [0304]. Jannard illustrates these components in Figure 22:

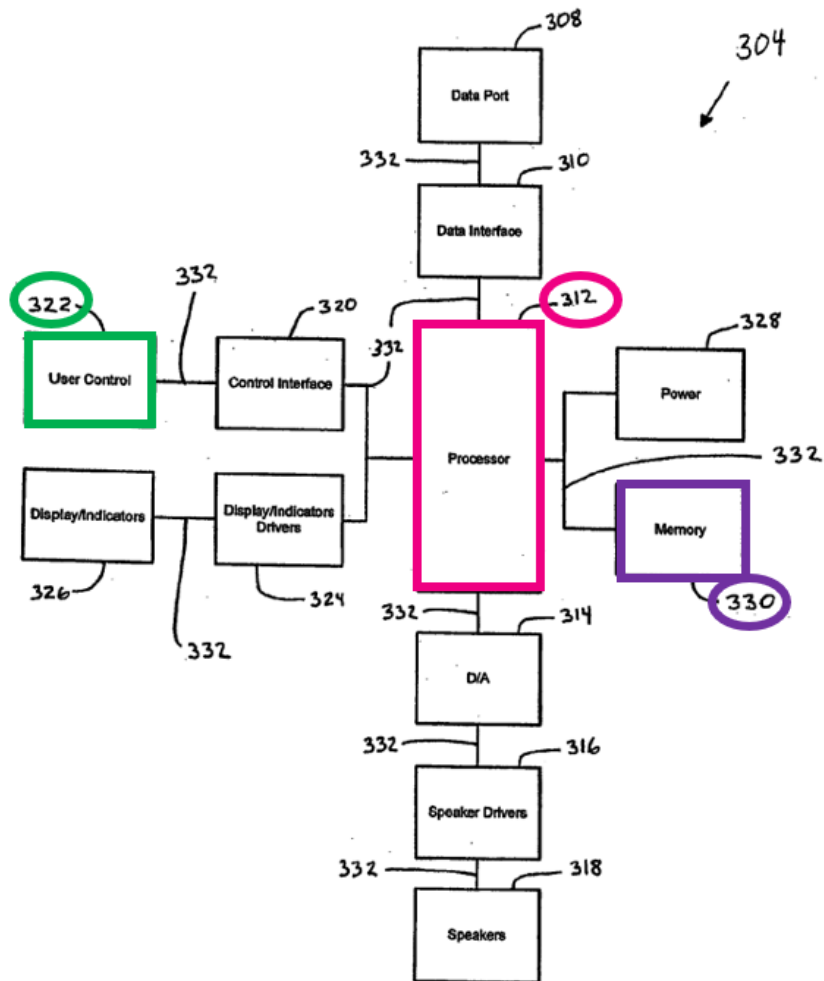


Figure 22

Jannard, Fig. 22

59. Jannard explains that the audio device 10 is “supported on the head” of a human, and support 12 is “supported by the head.” Jannard, [0072]-[0073]. The “support 12 can be in the form of a hat, sweatband, tiara, helmet, headphones, visor, and any of a variety of eyewear such as goggles, masks, face shields and eyeglasses.” Jannard, [0074]. Figures 11 and 12 illustrate an embodiment of these eyeglasses. As seen in Figures 11 and 12, the eyeglasses house “**microphone**

incoming phone call where the audio device 10D is used as an audio interface for a cellular phone.” Jannard, [0259]. Other than a button like **button 152**, the user controls 322 can include a “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.” Jannard, [0316]. Jannard describes that the “user controls 322 include a **microphone** that receives a voice command” from a user.” Jannard, [0316].

61. To enable the audio device to “accomplish complex electronic functions, while retaining a sleep, low weight configuration,” Jannard describes “off board[ing]” functionality to “source electronics.” Jannard, [0177]; *see also* Jannard [0179]. The “source electronics may include a computing device, such as a computer, server, a network, drive, RAM, ROM, or other non-removable or removable memory chip.” Jannard, [0191]. Jannard also offers a “cellular phone” as an exemplary source electronic. Jannard, [0297].

2. Overview of Wittenberg

62. Wittenberg discloses a “method and system for the remote processing of locally generated voice commands” “for controlling a local device” at a local site. Wittenberg, 2:35-40.

63. Common local device 160 includes a “**local command processor 110**” and “local controllable device 120” that a user wants to control. Wittenberg,

4:26-35. **Local command processor 110** “includes a microphone or similar mechanism for picking up acoustic signals” like voice commands and transmits the corresponding audio signal to remote processor 130. Wittenberg, 4:44-63, 5:32-45. “**Remote processor 130** includes at least a speech recognition processor and a voice recognition processor.” Wittenberg, 5:32-33. “Although not shown in FIG. 1, **remote processor 130** also includes a receiver capable of receiving audio signals from communication link 140 and a transmitter for transmitting resulting command signals over communication link 150 (e.g., a modem).” Wittenberg, 5:33-39. Once a voice command is “successfully identified via the speech recognition processing,” then “the **remote processor** transmits a command signal to local controllable device 120.” Wittenberg, 6:30-45.

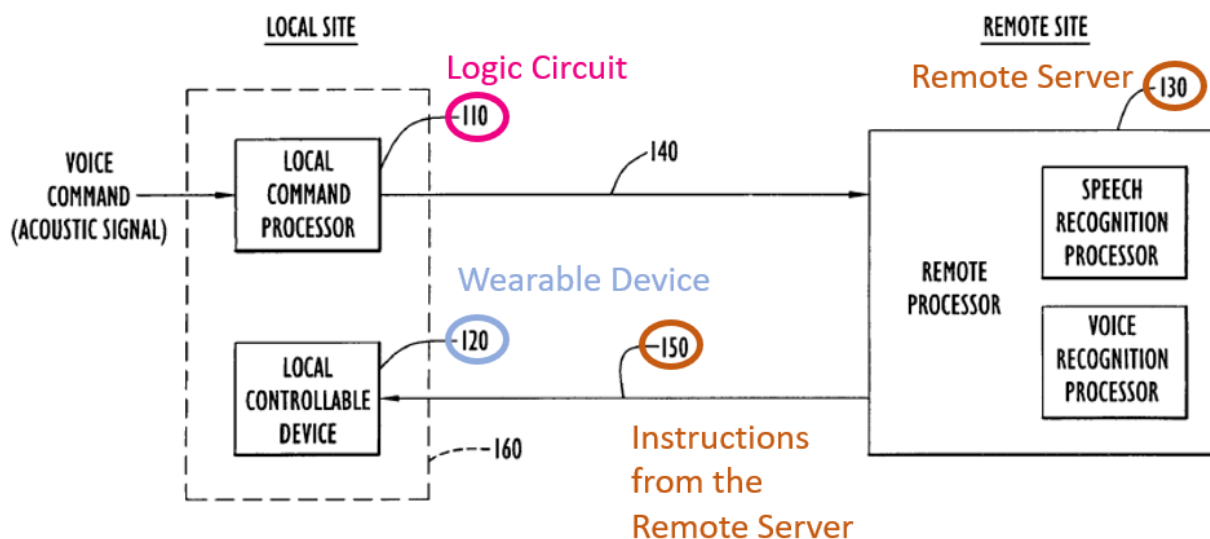


FIG.1

Wittenberg, Fig. 1

64. Wittenberg offers examples of applications of its invention in “a military or public-safety radio, a wireless telephone, a television, a computer, a security system for a premises or vehicle, etc.” Wittenberg, 2:40-43. Wittenberg also provides examples of different voice commands that a user could give to control the local device, including “switch to sub-net 3,” “order pay-per-view program 25,” “switch to channel 7,” “switch to vibrate only mode,” or “download Windows updates.” Wittenberg, 3:15-29.

3. Motivation to Apply Wittenberg’s Teachings to Jannard

65. Jannard and Wittenberg are analogous art. Jannard and Wittenberg are in the same field as the ’692 patent, which includes audio device control processing. E.g., ’692 patent, 2:16-29; Jannard, Abstract, [0013], [0289]; Wittenberg, 1:8-12, 2:35-43. Jannard and Wittenberg are reasonably pertinent to alleged problems addressed by the ’692 of combining the capacity of several devices into one convenient wearable audio device. *E.g.*, ’692 patent, 1:22-34 (“Present audio content playing devices are separated from the headphone system that normally contains the speakers (also referred to as receivers). The reason for this has typically been that audio content has been stored on disks that require a separate playing system. However, even with the advent of storing audio content on non-disk RAM (Random Access Memory) storage systems, the audio content player has been separated from the earpiece system (*e.g.*, plug in headphones or

earbuds). Combining the capacity for audio download and playing in an earpiece system is not obvious over related art since the user interaction system (*e.g.*, play button, keyboard system) does not readily appear compatible with the size of an earpiece device and the difficulty of user interaction.”); Jannard, Abstract, [0189] (“[A] wearable electronic interface 601 for providing data from one or more selected data sources to the wearer....Communication between the primary source 602 and any secondary source 604 and the interface 601 is accomplished...”); Wittenberg, 2:25-31 (“There are a numerous circumstances in the government/military, commercial, and private sectors where it would be advantageous to control a local device with a locally-generated voice command without placing the demands of speech processing on a local device and with the capability to provide a significant degree of security in the process of interpreting and carrying out voice commands”). For example, the ’692 patent’s goal is to operate “as simply as possible,” and both Jannard and Wittenberg address a similar goal. ’692 patent, 3:29. For example, Wittenberg describes “it would be advantageous to control a local device with a locally-generated voice command without placing the demands of speech processing on a local device.” Wittenberg, 2:25-29. Likewise, Jannard describes its system as avoiding the “bulk and weight” of carrying several different devices. Jannard, [0177].

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

apply Wittenberg's teachings of a remote server with speech processing and instruction transmission to a local device to Jannard's wearable device to advantageously "off load[]" functionality to another device, without "placing [] demands" on the local device. Wittenberg, 2:25-29; Jannard, [0177]. Both Jannard and Wittenberg expressly suggest and motivate this off loading. Jannard teaches that "off board[ing]" its functionality to "source electronics device," "such as...a server," will advantageously enable the wearable device to "accomplish complex electronic functions, while retaining a sleek, low weight configuration." Jannard, [0177], [0179], [0191]. And Wittenberg teaches advantageous remote processing of voice commands and resulting instruction transmissions to the local device "without placing the demands of speech processing on a local device." Wittenberg, 2:27-29. As such, off-loading allows the benefit of increasing functionality, while avoiding adding bulk and demand to a device.

67. Thus, in light of the above teachings, a person of ordinary skill in the art would have found it routine, straightforward and advantageous to apply Wittenberg's teachings of using a remote server to analyze audio signals for voice commands and send corresponding instructions to a local device to Jannard's teachings of a voice-command-responsive device. This would allow the device to maintain a sleek profile, while adding the benefit of increased functionality. In light of the above teachings, a person of ordinary skill in the art would have had a

reasonable expectation of success in applying Wittenberg’s teachings of remote server speech processing to Jannard’s teachings of a voice command-responsive device. Thus, a person of ordinary skill in the art would have known such a combination (yielding the claimed limitations) would predictably work and provide the expected functionality.

4. Invalidity of Claims 1-11, 13, 14, 16-20 Over Jannard in view of Wittenberg

68. As I discuss below, Claims 1-11, 13, 14, 16-20 are rendered obvious by Jannard in view of Wittenberg. In addition to my discussion in Sections I, IV, and V, the following discussion identifies example disclosures followed by bulleted excerpts that each disclose Claims 1-11, 13, 14, 16-20.

a. Element [1.pre]

69. **Element [1.pre]** recites “**A wearable device, comprising.**” This element is met by Jannard.

70. To the extent the preamble is limiting, Jannard discloses a wearable device (*e.g.*, “wearable audio device”/“audio device 10D”). *See, e.g.*, Jannard:

- Fig. 12

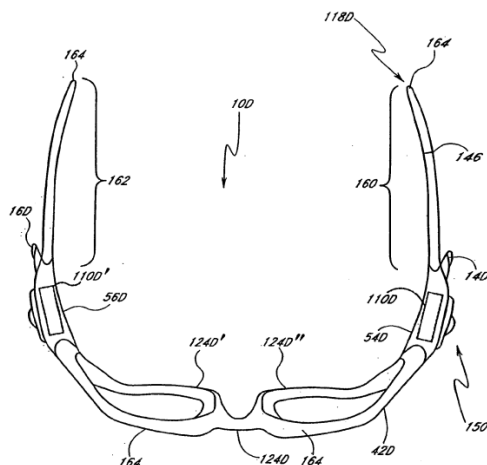


FIG. 12

- [0018]: “FIG. 1 is a schematic representation of a front elevational view of a **wearable audio device supported by a human head.**”
- [0241]: “With reference to FIGS. 10-12, a modification of the audio device 10C is illustrated therein and identified generally by the reference numeral 10D. The components of the audio device 10D which are the same as the components in the **audio devices 10, 10A, 10B, and 10C** are identified with the same reference numerals, except that a letter “D” has been added.”
- [0015]: “In accordance with a further aspect of the present invention, there is provided **a wearable electronically enabled interface system**, for providing audio and or video input to the wearer from at least two sources. The system comprises **a wearable support**, for carrying at least a portion of the interface.”
- [0020]: “FIG. 3A is a front, left side, and top perspective view of a modification of the **wearable audio device** illustrated in FIGS. 1 and 2.”
- [0073]-[0074]: “The **audio device 10** is illustrated as being supported on the head 18 of a human. The head 18 includes a nose 19, and left and right ears 20, 22. The schematic representation of human ears 20 and 22 are intended to represent the tissue forming the pinna of a human ear. With reference to FIG. 2, the meatus of the external auditory canal 24 is illustrated schematically as a circle (in phantom) generally at the center of the left ear 20. The support 12 is configured to be supported by the head 18. Thus, **the support 12 can be in the form of any of a variety of**

- wearable garments or devices, such as any known headwear.* For example, but without limitation, the support 12 can be in the form of a hat, sweatband, tiara, helmet, headphones, visor, and any of a variety of eyewear such as goggles, masks, face shields and eyeglasses.”
- [0089]: “The eyeglass 40 is also provided with a pair of generally rearwardly extending ear stems 54, 56 configured to retain the eyeglass 40 on the head of a wearer. In addition, an open region 58 is configured to receive the nose of the wearer, as is understood in the art. The open region 58 may optionally be provided with a nose piece, either connected to the lens orbitals 48, 50, or the bridge 52, or directly to the lenses, depending on the particular embodiment. Alternatively, the nose piece may be formed by appropriately sculpting the medial edges of the orbitals 48, 50 and the lower edge of the bridge 52, as in the illustrated embodiment.”
 - [0181]: “With reference to FIGS. 9A and 9B, the ***audio device 10C preferably comprises a wearable wireless audio interface device*** which includes a support 12C supported on the head 18 of a user by the support 26C and includes an interface device 110. The interface device 110 includes a power source 112, a transceiver 114, an interface 116, and an antenna 118.”
 - [0274]: “In this embodiment, additional functional attachments may be provided as desired anywhere along the length of the ***frame***, lenses or orbitals of the audio device 10E. For example, earphones may be directed towards the wearer's ear from one or two earphone supports extending rearwardly from the front of the ***eyeglass***, down from the top of the audio device 10E or forwardly from the rear of the audio device 10E. Similarly, one or more microphones may be directed at the wearer's mouth from one or two microphone supports connected to the orbitals or other portion of the audio device 10E.”

71. Jannard teaches a “wearable audio device supported by a human head.” Jannard, [0018]. The wearable audio device can take many forms, including eyeglasses, goggles, masks, and face shields. Jannard, [0074]. As figure 12 illustrates, the wearable audio device may be supported by an eyeglasses form.

b. Element [1.a]

72. **Element [1.a] recites “a housing for the wearable device.”** This element is met by Jannard.

73. Jannard discloses a housing (*e.g.*, “eyeglass”) for the wearable device (*e.g.*, “audio device 10D”). *See, e.g.*, Jannard:

- Fig. 12 (annotated)

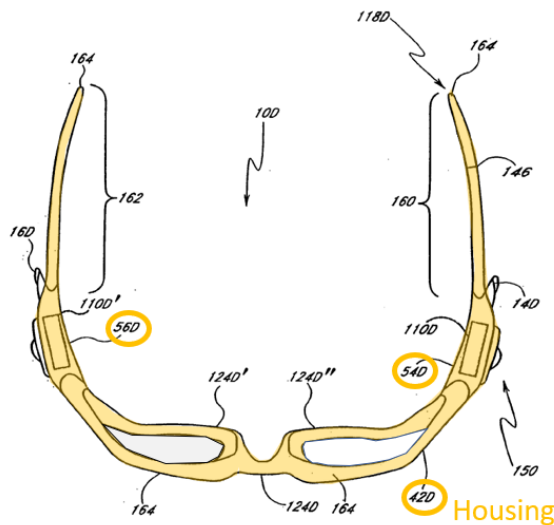


FIG. 12

- [0083]: “In the illustrated embodiment of the audio device 10A, the support 12A is in the form of ***an eyeglass 40***. The eyeglass 40 comprises a frame 42 which supports left and right lenses 44, 46.”
- [0241]: “***The components of the audio device 10D which are the same as the components in the audio devices 10, 10A, 10B, and 10C are identified with the same reference numerals, except that a letter “D” has been added.***”
- [0242]: “In the ***audio device 10D***, the microphone 124D can be disposed in the ***frame 42D***. In particular, ***the microphone 124D can be disposed in the bridge portion 52D.***”
- [0243]: “With reference to FIG. 12, the interface device 110D can be disposed in one of the ***ear stems 54D, 56D.***”

- [0255]: “Thus, where at least a portion of a frame of ***an eyeglass*** is used...”
- [0074]: “The support 12 is configured to be supported by the head 18. Thus, the support 12 can be in the form of any of a variety of wearable garments or devices, such as any known headwear. For example, but without limitation, the support 12 can be in the form of a hat, sweatband, tiara, helmet, headphones, visor, and any of a variety of eyewear such as goggles, masks, face shields and ***eyeglasses***.”

74. See my discussion for element [1.pre] above (§IX.A).

75. As shown in Figure 12, Jannard discloses “an eyeglass” that includes frame 42D and stems 54D, 56D. Jannard, [0241]-[0243], [0255]. The “eyeglass” houses components of audio device 10D, including interface 110D in ear stems 54D, 56D, and microphone 124D in frame 42D. Jannard, [0242]-[0243], Fig. 12.

c. Element [1.b]

76. **Element [1.b] recites “an ambient microphone within or on the housing.”** This element is met by Jannard.

77. Jannard discloses an ambient microphone (*e.g.*, “microphone 124D”; “banks of microphones” with “one bank providing an ambient noise-cancellation function”) within or on the housing (*e.g.*, within the “eyeglass”). *See, e.g.*, Jannard:

- Fig. 11 (annotated)

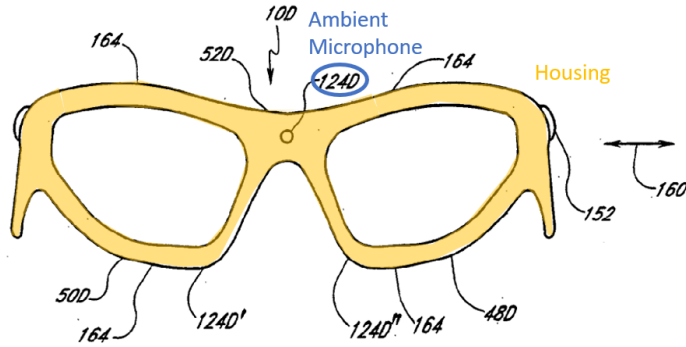


FIG. 11

- [0242]: “In the audio device 10D, the **microphone 124D can be disposed in the frame 42D**. In particular, the microphone 124D can be disposed in the bridge portion 52D.”
- [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. The microphone banks can be positioned at any suitable location or combination of locations (e.g., on the audio device, within the audio device, opposing sides of the audio device, or the like).** In one embodiment, automatic switching of the speaking-microphone and noise-canceling-microphone banks functions advantageously enhances ease of use. 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.”
- [0159]: “The audio device 10A ” is in the form of a eyeglass 12A ” having a frame 40A ” . **The audio device 10A ” also includes at least one microphone 75. Advantageously, the microphone 75 is disposed so as to face toward the wearer.**”

- [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.”
- [0242]: “In the audio device 10D, the **microphone 124D can be disposed in the frame 42D**. In particular, the **microphone 124D can be disposed in the bridge portion 52D**. Alternatively, the microphone 124D can be disposed along a lower edge of the right orbital 50D, this position being identified by the reference numeral 124D' . Further, the microphone could be positioned in a lower edge of the left orbital 48D, this position being identified by the reference numeral 124D'' . Optionally, two microphones can be disposed on the frame 42D at both the positions 124D' and 124D'' . Similarly to the microphone 75, the microphones 124D' , 124D'' **preferably are positioned so as to face toward the user**. Thus, the microphones 124D' , 124D'' can be protected from wind and noise. The microphones 124D, 124D' , 124D'' can also be constructed in accordance with any of the forms of the microphone 75 discussed above with reference to FIGS. 4A, 4B, 5A, 5B.”
- [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.”
- Fig. 4B (annotated)

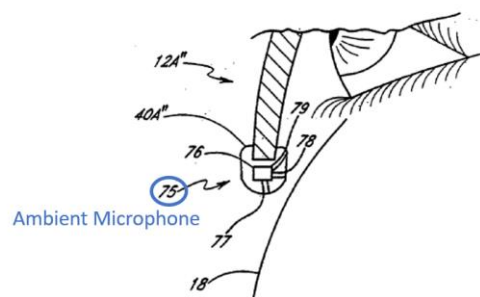


FIG. 4B

- Fig. 9A (annotated)

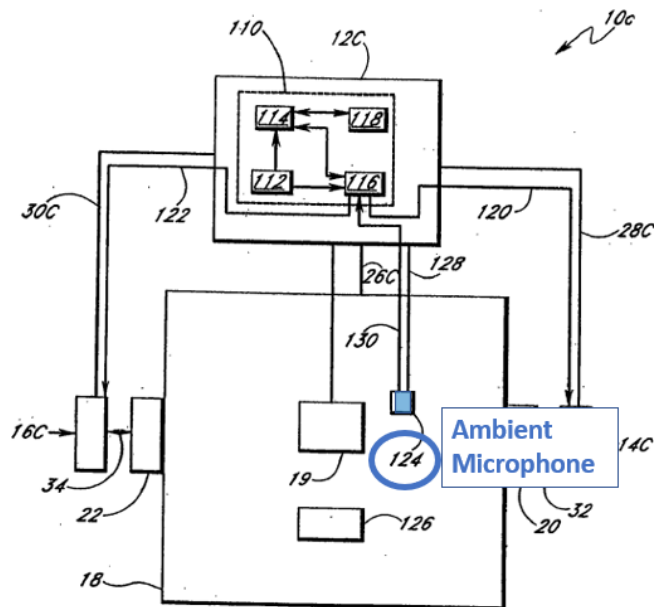


FIG. 9A

- Fig. 9B

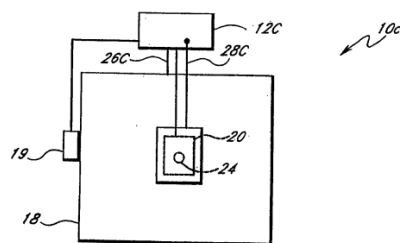


FIG. 9B

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

79. Jannard teaches “the microphone 124D can be disposed in the bridge portion 52D,” which is a part of the eyeglass support. Jannard, [0242]. Thus, microphone 124D is within or on the eyeglass in the bridge portion 52D of the eyeglass. Jannard, [0242]. Further, in Figure 11, Jannard illustrates microphone

124D as facing outward. Jannard, Fig. 11. Because microphone 124D faces outward, it picks up environmental noise, which is ambient noise.

80. As another example of Jannard's disclosure of an ambient microphone, Jannard discloses "at least two banks of microphones, with one bank acting as a speaker of received and one bank providing an ambient noise-cancellation function." Jannard, [0265]. The microphone banks "can be positioned at any suitable location or combination of locations (e.g., on the audio device, within the audio device, opposing sides of the audio device, or the like)." Jannard, [0265]. A microphone bank offering ambient noise cancellation function functions as an ambient microphone because the microphone must pick up ambient noise in order for the noise cancellation to determine what needs to be cancelled. And a person of ordinary skill in the art would have understood that Jannard's disclosure of microphones "within" or "on" the "audio device" to be within or on the eyeglass. At the very least, a person of ordinary skill in the art would have found it obvious for such microphones to pick up ambient noise with the microphone so that ambient noise can be cancelled. For example, the use of ambient microphones to support ambient noise cancellation had been well known for over a decade before the '692 patent's alleged priority date. *See, e.g.*, Ex, 1016 (U.S. 5,251,263 ("Adaptive noise cancellation and speech enhancement system and apparatus therefor"), dated 1993).

d. Element [1.c]

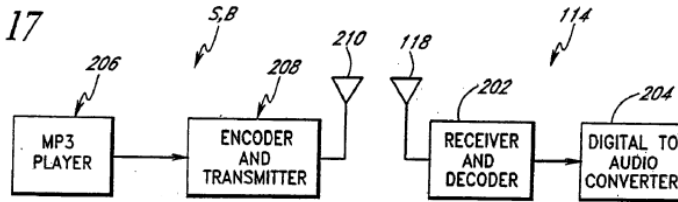
81. Element [1.c] recites “a communication module within or on the housing.” This element is met by Jannard.

82. Jannard discloses a communication module (e.g., “transceiver 114”) within or on the housing (e.g., “in the left ear stem 54D” within the “eyeglass”).

See, e.g., Jannard:

- [0244]: “In one arrangement, ***the transceiver 114, interface 116, and the antenna 118 can be disposed in the left ear stem 54D*** with the battery 112 being disposed in the right ear stem 56D.”
- [0012]: “The ***wearable cell phone and music system may additionally comprise radiofrequency electronics such as a Bluetooth receiver or transceiver***, for wireless communication between the cell phone and the support.”
- [0183]: “***The transceiver 114 can be in the form of a digital wireless transceiver for one-way or two-way communication.*** For example, the transceiver 114 can be a transceiver used in known wireless networking devices that operate under the standards of 802.11a, 802.11b, or preferably, the standard that has become known as BLUETOOTH™. As illustrated in BLUETOOTH™-related publications discussed below, the BLUETOOTH™ standard advantageously provides low-cost, low-power, and wireless links using a short-range, radio-based technology. Systems that employ the BLUETOOTH™ standard and similar systems advantageously allow creation of a short-range, wireless “personal area network” by using small radio transmitters. Consequently, with BLUETOOTH™-enabled systems and similar systems, components within these systems may communicate wirelessly via a personal area network.”
- Fig. 17

FIG. 17



- [0016]: “The wearable support may comprise an eyeglass frame. The data port may comprise a radiofrequency receiver or **transceiver**. The interface may comprise at least one speaker, and, in certain implementations, the interface comprises two speakers. The interface may additionally comprise at least one microphone. The interface may further comprise a video display. The selector may comprise a wearer activated control such as a button, switch or voice activated electronic control.”
- [0100]: “The audio device 10A can be used as an audio output device for any type of device which provides an audio output signal. The audio device 10A can include an audio input terminal or jack disposed anywhere on the eyeglass 40 for receiving a digital or analog audio signal. Preferably, wires connecting the input jack (not shown) with the speakers 14A, 16A extend through the interior of the ear stems 54, 56 so as to preserve the outer appearance of the eyeglass 40. Alternatively, the audio device 10A can include a wireless receiver or **transceiver for receiving digital signals from another device.**”
- [0172]: “With reference to FIGS. 5, 6 and 7, the support arms 100, 102 extend downwardly from the stems 96, 98, respectively. In the illustrated embodiment, the support arms 100, 102 extend in an “L” shape. In particular, the support arm 100 extends from the stem 96 to a point just forward (anterior) from the tragus of the user's ear 20. From this point, the support arm 100 extends rearwardly so as to support the speaker 14B at a position juxtaposed and spaced from the ear 20. Preferably, the speaker 14B is maintained in a position from about 2 mm to 3 cm from the tragus of the ear 20. Similarly to the audio device 10A, **the audio device 10B can include an audio input through a wired arrangement or through a wireless transceiver.**”
- [0181]: “With reference to FIGS. 9A and 9B, the audio device 10C preferably comprises a wearable wireless audio interface device which includes a support 12C supported on the head 18 of a user by the support

26C and includes an interface device 110. The interface device 110 includes a power source 112, **a transceiver 114**, an interface 116, and an antenna 118.”

- [0183]: “The **transceiver 114 can be in the form of a digital wireless transceiver for one-way or two-way communication**. For example, the **transceiver 114 can be a transceiver used in known wireless networking devices that operate under the standards of 802.11a, 802.11b**, or preferably, the standard that has become known as BLUETOOTH™. As illustrated in BLUETOOTH™-related publications discussed below, the BLUETOOTH™ standard advantageously provides low-cost, low-power, and wireless links using a short-range, radio-based technology. Systems that employ the BLUETOOTH™ standard and similar systems advantageously allow creation of a short-range, wireless “personal area network” by using small radio transmitters. Consequently, with BLUETOOTH™-enabled systems and similar systems, components within these systems may communicate wirelessly via a personal area network. Personal area networks advantageously may include voice/data, may include voice over data, may include digital and analogue communication, and may provide wireless connectivity to source electronics. Personal area networks may advantageously have a range of about 30 feet; however, longer or shorter ranges are possible. The antenna 118 can be in the form of an onboard antenna integral with the transceiver 114 or an antenna external to the transceiver 114. In another implementation, the **transceiver 114 can support data speeds of up to 721 kilo-bits per second as well as three voice channels**.”
- [0184]: “In one implementation, **the transceiver 114 can operate at least two power levels**: a lower power level that covers a range of about ten yards and a higher power level. The higher level covers a range of about one hundred yards, can function even in very noisy radio environments, and can be audible under severe conditions. The transceiver 114 can advantageously limit its output with reference to system requirements. For example, without limitation, if the source electronics B is only a short distance from audio device 10C, the transceiver 114 modifies its signal to be suitable for the distance. In another implementation, the transceiver 114 can switch to a low-power mode when traffic volume becomes low or stops.”

- [0185]: “The interface 116 can be configured to **receive signals from the transceiver 114 that are in the form of digital or analog audio signals**. The interface 116 can then send the audio signals to the speakers 14C, 16C through speaker lines 120, 122, respectively, discussed in greater detail below.”
- [0187]: “The microphone 124 is connected to the interface 116 through a microphone line 130. Thus, **the transceiver 114 can receive audio signals from the microphone 124 through the interface 116**. As such, the 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. The microphone 124 can also be in any of the forms discussed above with reference to the microphone 75.”
- [0211]: “The data port 608 is any device capable of receiving information from a primary source 602 (or secondary source 604) via its associated communication link 606. For example, in one embodiment, the data port 608 is a physical connector such as a mini-USB connector depending upon the nature of the communication link 606. In such embodiment, the primary source 602 or secondary source 604 might be coupled to the wearable electronic interface 601 via a USB cable having a mating mini-USB connector on at least one of its ends. In another embodiment, the data port 608 includes a wireless transceiver for providing wireless communication between the primary source 602 (or secondary source 604) and the wearable electronic interface 601. For example, in one embodiment, the data port 608 includes a BLUETOOTH receiver or **transceiver**. The data port 608 includes any data port 308 described in greater detail below with respect to FIG. 22.
- Fig.12 (annotated to show where transceiver 114 is located)

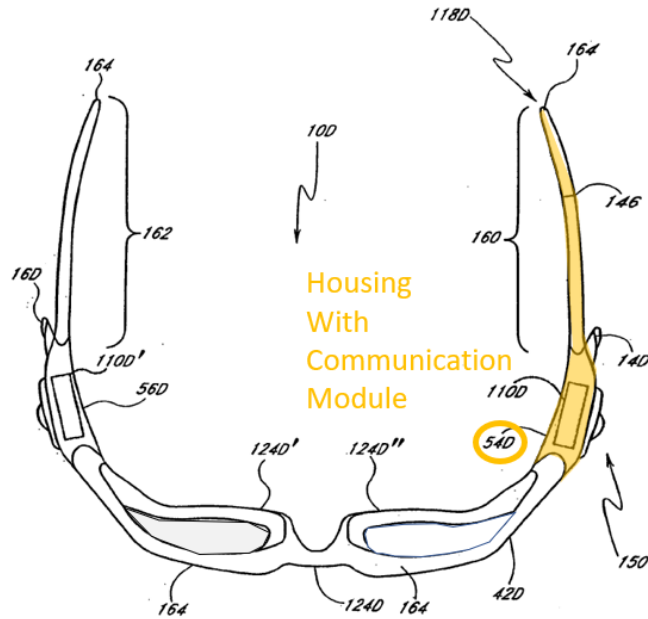


FIG. 12

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

84. Jannard teaches “transceiver 114 ... in the form of a wireless transceiver for one-way or two-way communication.” Jannard, [0183]. While the transceiver 114 is not labeled in Figure 12, Jannard teaches that “transceiver 114...can be disposed in the left ear stem 54D” of the eyeglass, which can be seen (and labeled above) in Figure 12. Jannard, [0244], see Fig. 12.

e. Element [1.d]

85. **Element [1.d] recites “a logic circuit communicatively coupled to the ambient microphone.”** This element is met by Jannard.

86. Jannard discloses a logic circuit (e.g., “processor 312”) communicatively coupled to the ambient microphone (e.g., “receive commands from a user via the user controls 322,” which “include a microphone that receives a

voice command”). See, e.g., Jannard:

- Fig. 22

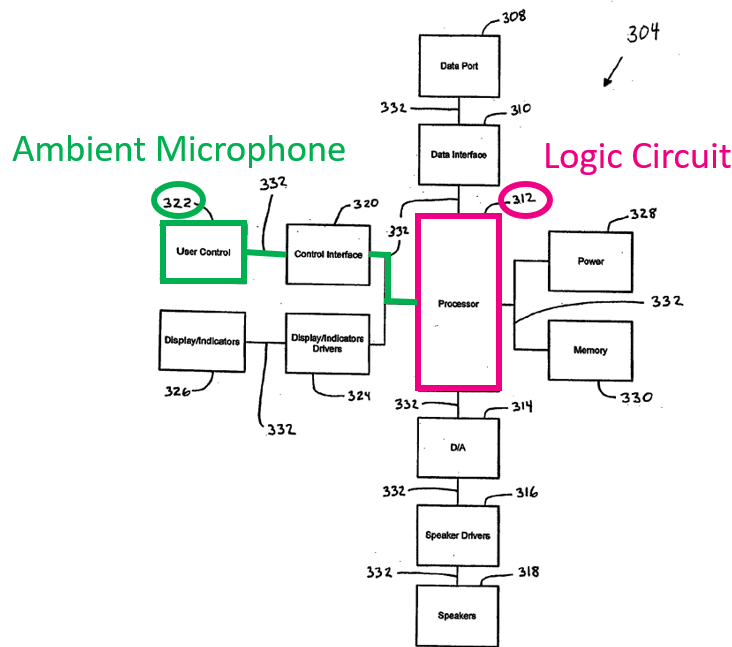


Figure 22

- [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.**”)
- [0316] (“**In one embodiment, the user controls 322 include a microphone that receives a voice command.**”)
- [0265] (“[T]he 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.”)
- [0312] (“The audio **device** 304 may be any of the audio devices described above with respect to FIGS. 1-19....”).

- [0242] (“In the audio device 10D, the **microphone 124D** can be disposed in the frame **42D**. In particular, the microphone 124D can be disposed in the bridge portion 52D. Alternatively, the microphone 124D can be disposed along a lower edge of the right orbital 50D, this position being identified by the reference numeral 124D’. Further, the microphone could be positioned in a lower edge of the left orbital 48D, this position being identified by the reference numeral 124D”. Optionally, two microphones can be disposed on the frame 42D at both the positions 124D’ and 124D”. Similarly to the microphone 75, the microphones 124D’, 124D” preferably are positioned so as to face toward the user. Thus, the microphones 124D’, 124D” can be protected from wind and noise. The microphones 124D, 124D’, 124D” can also be constructed in accordance with any of the forms of the microphone 75 discussed above with reference to FIGS. 4A, 4B, 5A, 5B.”)
- [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 play back the MP3 files through the speakers 14A’, 16A’. Suitable electronics for enabling and amplifying MP3 storage and playback are well known in the art, and may be commercially available from Sigmatel, Inc. or Atmel, Inc. Thus, further description of the hardware and software for operating the device 286 as a storage and playback device is not necessary for one of ordinary skill in the art to make and use the inventions disclosed herein.”
- [0214]: “The selector 610 may include a mechanical, electrical, or an electro-mechanical switch. For example, in one embodiment, the selector 610 includes a manually operable control such as a toggle switch, a rocker switch, a jumper, a dial, a button, a knob, or a combination thereof. In another embodiment, the selector 610 includes an electronically operable control such as a transistor, a bank of transistors, a relay, a circuit, logic, a RAM, a ROM, a PLD, an EPROM, an EEPROM, a microprocessor, a microcontroller, a multiplexor, a demultiplexor, or a combination thereof. 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.***”

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

88. Jannard teaches “processor 312 [] receive[s] commands from a user via the user controls 322.” Jannard, [0315]. Jannard goes on to teach that “user controls 322 include a microphone that receives a voice command.” Jannard, [0316]. Therefore, in audio device 10D, microphone 124D receives the voice command. Jannard, [0242]. Similarly, when Jannard’s microphone bank is used, the voice command is received via the microphone bank with “voice recognition” and provided to processor 312. Jannard, [0265].

89. Since Jannard teaches that the user controls 322 include a microphone that receives voice controls, a person of ordinary skill in the art would have understood the processor 312 is communicatively coupled to the microphone as illustrated in Figure 22, in order to receive these commands and to pass them along to the server so they can be acted on (see my discussion for element [1.g], below). See Jannard, Fig. 22. In particular, a POSITA would have understood that any input device (such as a microphone, button, etc.) would have the connections illustrated in Figure 22, including a connection to the processor 312. At the very least, it would have been obvious for the processor 32 to be communicatively coupled to the microphone of user controls 322 for the same reason.

f. Element [1.e]

- [0321]: “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.”
- [0322]: “**The memory module 330 may be permanently contained within the audio device 304**, or may be removable.”
- [0134]: “The internal cavity 282, in the illustrated embodiment, is configured to receive electronics such as a printed circuit board 284. In the illustrated embodiment, the printed circuit board 284 includes one switch for each of the buttons 73 c, 73 d, and 73 e. Additionally, the printed circuit board 284 includes **an audio file storage** and playback device 286.”

92. See my discussion for element [1.d] above (§IX.A).

93. Jannard teaches “memory module 330, which...stores audio data files.” Jannard, [0321]. Jannard also teaches a “memory... and ... processor ... configured to operate together.” Jannard [0135]. Jannard shows the memory and processor communicatively coupled in Fig. 22. Jannard [0135]. For example, 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. Jannard, [0307]. 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.1015 at 365.

94. At the very least, a person of ordinary skill in the art would have found it obvious to include instructions for processor 312 in memory coupled to processor 312. This is because a processor runs on instructions, and such instructions are stored in memory so as to allow the processor to access and operate on those instructions. A person of ordinary skill in the art would have recognized that the alternative to storing the instructions in memory in the audio device, for example storing the instructions 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. This would also require additional coordination between the audio device and remote memory, and further complicate the system.

g. Element [1.f]

95. **Element [1.f] recites “an interaction element, the interaction element and logic circuit cooperatively initiate control of media content from the wearable device or initiate operations of the wearable device.”** This element is met by Jannard.

96. Jannard discloses an interaction element, the interaction element (*e.g.*, “user controls 322” including a “button, dial ..., microphone, and/or any other input device that may be used by a user to control the audio device”) and logic circuit (*e.g.*, “processor 312” see [1.d]) cooperatively initiate control of media

content (e.g., “processor 312 receive[s] commands from a user via the user controls 322” to “perform any one or a combination of various functions with respect to an audio data file”) from the wearable device (e.g., “download[ed] from the content source...to the audio device”) or initiate operations of the wearable device (e.g., “button 152 could be used to answer an incoming phone call”). See, e.g., Jannard:

- Fig. 22 (annotated)

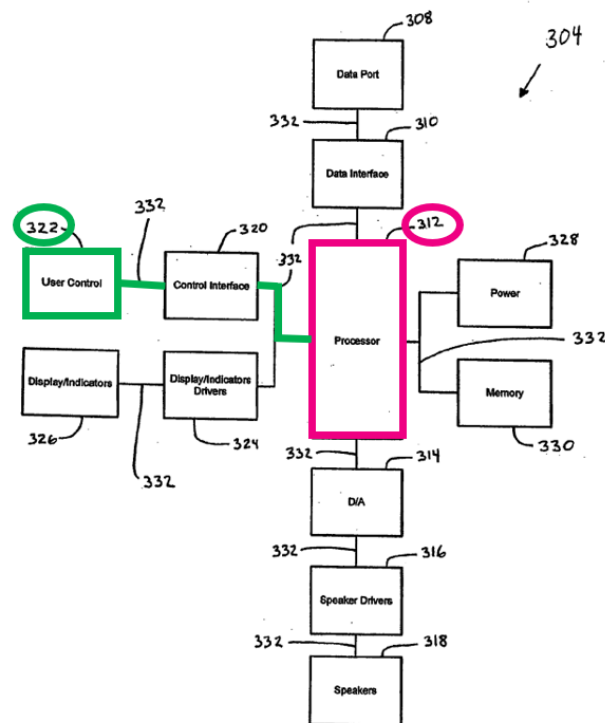


Figure 22

- [0257]: “Preferably, the audio device 10D includes a user interface device 150 configured to transmit user input signals to the interface 116 and/or the transceiver 114. In the illustrated embodiment, ***the user interface device 150 is in the form of a 3-way button. The 3-way button 152 is configured to have three modes of operation.***”
- Fig. 11 (annotated)

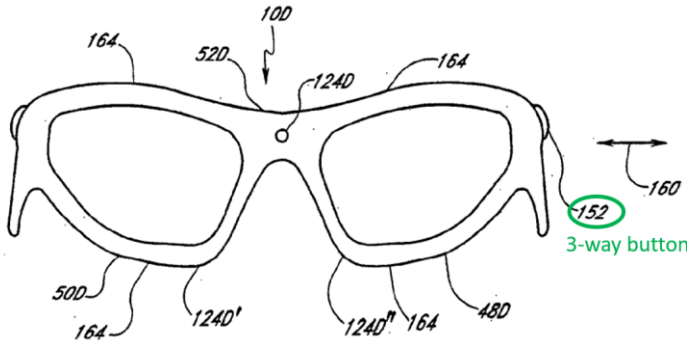


FIG. 11

- [0258]: “***In one illustrative and non-limiting embodiment, the button 152 can be used to control volume.***”
- [0259]: “In a further illustrative and non-limiting example, the medial-lateral movement of ***the button 152***, along the directions identified by the arrow 160, ***can be used to choose different functions performed by the transceiver 114 or the interface 116.*** For example, an inward movement of ***the button 152 could be used to answer an incoming phone call*** where the audio device 10D is used as an audio interface for a cellular phone.”
- [0302]: “The audio device 304 may be any of the audio devices described above with respect to FIGS. 1-19....”
- [0303]: “***In one embodiment, the audio network 300 is configured for the downloading of music from the content source 302 (e.g., a use’s computer) to the audio device 304.*** In another embodiment, the audio network 300 is configured for the uploading of content stored within the audio device 304 to the content source 302.”
- [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.”
- [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.***”

- [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.”
- [0016]: “The wearable support may comprise an eyeglass frame. The data port may comprise a radiofrequency receiver or transceiver. The interface may comprise at least one speaker, and, in certain implementations, the interface comprises two speakers. The interface may additionally comprise at least one microphone. The interface may further comprise a video display. The selector may comprise a wearer activated control such as **a button, switch or voice activated electronic control.**”
- [0122]: “As noted above, an aspect of one of the inventions disclosed herein includes the realization that where an electronic device that is worn in the same manner as a pair of eyeglasses includes **a user operable switch for controlling a function of the electronics**, the comfort of the wearer of the audio device can be enhanced where the switches are operable without transferring a substantial load to the head of the wearer. For example, where the electronic device includes buttons for controlling an aspect of the device, a further advantage is provided where a support surface is provided opposite the button such that a user can apply a balancing force to the actuation force applied to the button, thereby preventing a substantial force from being transferred to the head of the wearer.”
- [0123]: “With reference to FIG. 3I, the audio device 10A' can include at least one button 73 a. In the illustrated embodiment, the audio device 10A' includes **five buttons**; a first button 73 a and a second button 73 b mounted to the left ear stem 54', and a third button 73 c, a fourth button 73 d, and a fifth button 73 e mounted to the right ear stem 56'. Of course, this is one preferred embodiment of the arrangement of the buttons 73 a, 73 b, 73 c, 73 d, 73 e. Other **numbers of buttons and other arrangements of buttons are also applicable.**”
- [0124]: “With reference to FIG. 3I, the audio device 10A' can include at least one button 73 a. In the illustrated embodiment, the audio device 10A' includes **five buttons**; a first button 73 a and a second button 73 b mounted to the left ear stem 54', and a third button 73 c, a fourth button 73 d, and a fifth button 73 e mounted to the right ear stem 56'. Of course, this is one preferred embodiment of the arrangement of the

buttons 73 a, 73 b, 73 c, 73 d, 73 e. Other ***numbers of buttons and other arrangements of buttons are also applicable.***”

- [0125] “This provides a further advantage in that a repeated application of a force against the audio device 10A’ that is transferred to the head of the wearer of the audio device 10A’ is avoided. For example, where the audio 10A’ is in the form of eyeglasses 12A’, a wearer of the eyeglasses 12A’ can be subjected to irritations if the wearer repeatedly presses the eyeglasses 12A’ to actuate a switch. Further, such repeated loads can cause headaches. Thus, by configuring the ear stems 54A’ such that the ***button 73 a can be depressed*** without transferring a substantial load to the wearer of the ear glasses 12A’, such irritations and headaches can be avoided.”
- [0126]: “Further, by disposing the button 73 a on an upper portion of the ear stems 54A’, and by providing the ear stems 54A’ with an opposite lower surface that faces an opposite direction relative to the upper surface, a wearer can grasp the ear stems 54A’ from the side, as illustrated in FIG. 38, thereby allowing the user to counteract the actuation force required to actuate the button 73 a without having to insert a finger between a side of the wearer's head and ear stems 54A’. In any of the embodiments herein, the surface which opposes the buttons may be provided with any of a variety of tactile feedback structures, such as ridges or bumps, that have a predetermined alignment with respect to the buttons. This can assist the user in positioning their thumb in the identical position each time, so that the user, after a learning period, can rapidly reach for the controls, position their hand and identify with which ***button*** their fingers are aligned. See, for example, ***the tactile indicium illustrated in FIG. 24.***”
- [0127]: “FIG. 3J illustrates an exploded view of an exemplary embodiment of the audio device 10A’. As shown in FIG. 3J, the left side ear stem 54A’ defines an electronic housing portion 250 which defines an internal cavity 252 configured to receive electronic components. The electronics housing 250 includes an upper surface 254 and lower surface 260. The upper surface 254 extends generally outwardly from the ear stems 54A’ and around the internal cavity 252. The upper surface also includes apertures 256, 258 through which ***buttons 73 a, 73 b***, respectively, extend.”

- [0135]: “FIG. 3J illustrates an exploded view of an exemplary embodiment of the audio device 10A'. As shown in FIG. 3J, the left side ear stem 54A' defines an electronic housing portion 250 which defines an internal cavity 252 configured to receive electronic components. The electronics housing 250 includes an upper surface 254 and lower surface 260. The upper surface 254 extends generally outwardly from the ear stems 54A' and around the internal cavity 252. The upper surface also includes apertures 256, 258 through which **buttons 73 a, 73 b**, respectively, extend.
- [0137]: “A door 290 is configured to open and close the aperture through which the data port 288 is exposed. Preferably, the door 290 is hingedly engaged to the housing 280, in an identical or similar manner as the door 276. In the illustrated embodiment, the door 290 can be pivoted relative to housing 280, thereby exposing the **data transfer port 288**. In the illustrated embodiment, the data transfer port is configured to operate according to the universal serial bus (USB) transfer protocol. In one implementation of the invention, the earstem is provided with a mini USB port. The mini USB port enables both **downloading of digital music from a source into the eyeglass**, as well as charging a rechargeable battery carried by the eyeglass. Optical data ports may alternatively be used. As a further alternative, **MP3 files may be uploaded from a source using wireless systems**, such as BLUETOOTH® protocols, as is discussed below. Further, **the device 286 is configured to receive audio files from another computer, through the data transfer port 288 and to store the files into the memory incorporated into the device 286**.
- [0151]: “The **buttons 73 c, 73 d, 73 e** and the data transfer port 288 are connected to the device 286 through printed conduits incorporated into the **printed circuit board 284**.
- [0220]: “In one embodiment, during operation, the wearer of the wearable electronic interface 601 manually selects which input source 602, 604 is placed in communication with the interface output. The **wearer can switch input sources by activation of the selector at any time**. In another embodiment, the wearable electronic interface 601 automatically selects the particular input source 602, 604 for communication based upon a prioritization schedule configured by the wearer.

- [0291]: “With reference to FIG. 9A, ***the interface electronics 116 can include audio or aural menus that can be selected by user.*** 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. The voice signal can indicate that a first menu option is available. For example, but without limitation, the first menu choice can be to initiate a call. ***Thus, when the user pushes the button 150 the first time, the user will hear the words “initiate a call,” emanating from the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16 E. 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).***
- [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.

97. See my discussion for elements [1.pre] and [1.d] above (§IX.A).

98. Jannard teaches “processor 312 to receive commands from a user via the user controls 322.” Jannard, [0315]. As I described in element [1.d] above, “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.” Jannard, [0316]. “[U]ser controls 322 are used to perform any one or a combination of various functions with respect to an audio data file.” Jannard, [0317]. Jannard provides examples of the various functions, including “play, pause...rewind.... repeat playback... or perform any other function related to an audio data file.” Jannard, [0317]. Jannard further describes that “audio network 300 is configured for the downloading of music from the content source 302 (e.g., a user's computer) to the audio device 304.” Jannard, [0303]. Thus, processor 312 receives commands from the user via user controls 322, to perform various functions regarding the audio data files downloaded onto the audio device from a content source. Further, “button 152 ([an element of user controls 322]) could be used to answer an incoming phone call where the audio device 10D is used as an audio interface for a cellular phone.” Jannard, [0259].

h. Element [1.g]

99. **Element [1.g] recites “wherein the logic circuit is configured to cooperatively initiate control of media content of the wearable device in response to analysis of audio captured at the wearable device at the ambient microphone and analyzed at a remote server in response to instructions received from the remote server.”** This element is met by Jannard in view of Wittenberg.

100. Jannard discloses the logic circuit is configured to cooperatively

initiate control of media content of the wearable device (*e.g.*, see [1.d], [1.f]) in response to analysis of audio captured at the wearable device at the ambient microphone (*e.g.* “commands from a user via the user controls 322,” which “include a microphone that receives a voice command”). See, *e.g.*, Jannard:

- [0177] (“In general, the component configuration of FIG. 8 enables the audio device 10C to carry interface electronics with the user, such as audio output and audio input. However, ***the source electronics*** such as the MP3 player, cellular phone, computer or the like ***may be off board, or located remotely from the audio device 10C. This enables the audio device 10C to accomplish complex electronic functions, while retaining a sleek, low weight configuration....*** This accomplishes the function of ***off loading the bulk and weight of the source electronics from the headset.***”)
- [0257] Preferably, the audio device 10D includes a ***user interface device 150*** configured ***to transmit user input signals to the interface 116 and/or the transceiver 114.***”)
- [0303] (“***In one embodiment, the audio network 300 is configured for the downloading of music from the content source 302 (e.g., a user's computer) to the audio device 304.***”)
- [0307] (“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.”)
- [0316] (“In one embodiment, the user controls 322 include ***a...microphone...that may be used by a user to control the audio device 304....***[T]he user controls 322 include ***a microphone that receives a voice command.***”)
- FIG. 21

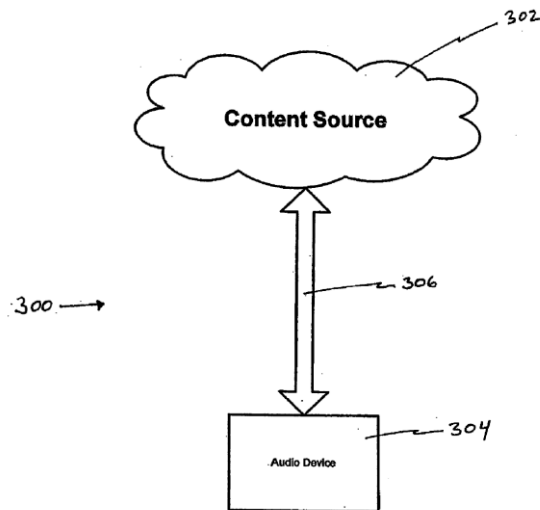


Figure 21

- [0011]: “There is provided in accordance with one aspect of the present invention, a wearable cell phone and music system. The system comprises a wearable support, having a cell phone in communication with the support and a music source in communication with the support. An electronic interface is carried by the support. A selector is provided, for regulating the transmission of data from at least one of the cell phone and the music source to the interface. The selector may be a button carried by the support (e.g., on an earstem or on an orbital in an eyeglass configuration) to allow the wearer to switch between the cell phone and music source.”
- [0012]: “The wearable cell phone and music system may additionally comprise radiofrequency electronics such as a Bluetooth receiver or transceiver, for wireless communication between the cell phone and the support. The system may additionally comprise radiofrequency electronics for wireless communication between the music source and the support. The music source may either be carried by the support, or exist remotely from the support and communicate wirelessly with the support. In one embodiment, the music source comprises an MP3 player.”
- [0016]: “The wearable support may comprise an eyeglass frame. The data port may comprise a radiofrequency receiver or transceiver. The interface may comprise at least one speaker, and, in certain implementations, the

interface comprises two speakers. The interface may additionally comprise at least one microphone. The interface may further comprise a video display. The selector may comprise a **wearer activated control such as a button, switch or voice activated electronic control.**”

- [0168]: “The audio device 10A ” can include electrical conduits extending through the frame 40A ” to an audio output jack (not shown). The audio output jack can be disposed at the end of the ear stems 54A ” , 56A ” , or anywhere else on the frame 40A ” . **Thus, a user can wear the audio device 10A ’ and use the microphone 75 in order to transform the voice of the wearer or other sounds into an electrical signal.** The electrical signal can be transmitted to another audio device, such as a palm top computer, a laptop computer, a digital or analog audio recorder, a cell phone, and the like. Additionally, the audio device 10A ” can include speakers, such as the speakers 14A ” , 16A ” illustrated in FIG. 3A. As such, the audio device 10A ” can be configured to provide two-way audio for the wearer, e.g., audio input being transmitted to the user through the speakers 14A ” , 16A ” , and audio output being transmitted from the wearer, through the microphone 75, and out through the audio output jack. As such, a user can use the audio device 10A ” for two-way audio communication in a comfortable manner.”
- [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.”

- [0187]: “The microphone 124 is connected to the interface 116 through a microphone line 130. Thus, the transceiver 114 can receive audio signals from the microphone 124 through the interface 116. As such, the 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. The microphone 124 can also be in any of the forms discussed above with reference to the microphone 75.”
- [0189]: “One embodiment of an input data management system 600 in accordance with the present invention is schematically illustrated in FIG. 8A. The data management system 600 includes a wearable electronic interface 601 for providing data from one or more selected data sources to the wearer. The interface 601 is in communication with a primary data source 602 and optionally at least one secondary source 604. Communication between the primary source 602 and any secondary source 604 and the interface 601 is accomplished via at least one communication link 606. In one embodiment, the wearable electronic interface 601 is in communication with one, two, three, or n secondary sources 604.”
- [0190]: “In general, the data input management system is configured to allow a 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. This allows the wearer to obtain the benefits of multiple input sources, without the need to replace or make any changes to the interface. As will be discussed in greater detail below, the user may select only a single data source for connection to the interface. Alternatively, the user may select one source as a primary input source and a second source as the secondary input source. The interface 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 system.”
- [0191]: “The primary source 602 and secondary source 604 may be any source, conduit, or provider of audio, video or audio/video information selected by the wearer or by the manufacturer. The examples identified below will therefore be designated generically as source electronics. For example, the source electronics may include a computing device, such as a computer, a server, a network, drive, RAM, ROM or other non-removable or removable memory chip.”

- [0236]: “In use, the **wearer may be listening to music from the primary source when an incoming phone call is received**. If the incoming call is from a number which has been preprogrammed into the system, the override function allows the name of the caller or other information associated with that incoming phone number to be presented audibly over the speakers. 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.
- [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.”

101. See my discussion for elements [1.b], [1.d], and [1.f] above (§IX.A).

102. As I explained in elements [1.b], [1.d] and [1.f], Jannard teaches

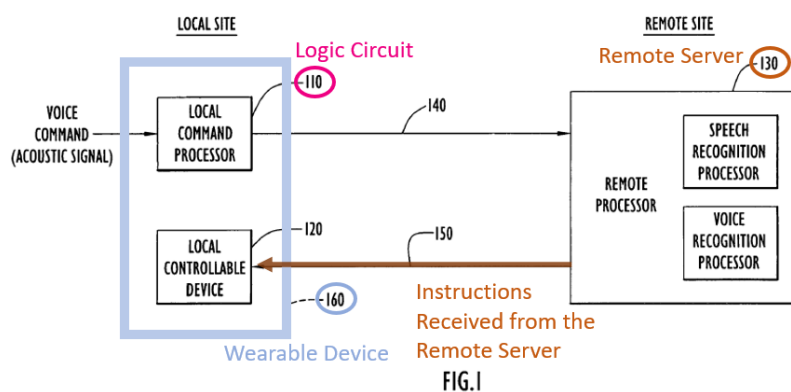
media content of audio device 10D is controlled by processor 312 through user controls 322, which include a microphone that receives voice controls. The voice controls are in the form of “voice commands” captured by the microphone.

Jannard, [0316]. Jannard further teaches “off-load[ing]” functionality from the audio device to “source electronics device,” “such as...a server” with “a speech recognition engine...that is properly trained to recognize the voice of the user.”

Jannard, [0177], [0179], [0191], [0297].

103. **Wittenberg discloses a logic circuit** (*e.g.*, “local command processor 110”) **cooperatively initiating control of a device** (*e.g.*, “alter[ing] a state of local controllable device 120” within “device 160”) **in response to analysis of audio analyzed at a remote server** (*e.g.*, “remote processor 130” “attempts to identify a voice command with the audio signal”) **in response to instructions received from the remote server** (*e.g.*, “local device 160” receives “a command signal” from “remote processor 130”). *See, e.g.*, Wittenberg:

- Fig. 1



- [Abstract] (“A technique for remotely processing a local audio command to control a local device includes: receiving at a local site an acoustic signal and generating a corresponding audio signal; transmitting the audio signal to a remote site; performing speech recognition processing on the audio signal at the remote site to determine whether the audio signal includes a command; performing voice recognition processing on the audio signal at the remote site to determine whether the audio signal has been supplied by an authorized user; generating a command signal in response to the audio signal including a command and being supplied by an authorized user; and transmitting the command signal to a device at the local site to effect a change in a state of the local device.”) Ex. 1006 (Wittenberg, Application as-filed, pg 22)
- 2:35-64 (“According to the present invention, there is provided a method and system for the **remote processing of locally generated voice commands**, and more particularly, **for controlling a local device operated by an operator at a local site...**The **operator controls the local device by speaking voice commands**. The **voice command** constitutes an acoustic signal that travels through the air and is **received by a microphone of a local command processor**. The **local command processor generates a corresponding audio signal** (e.g., an electric or electromagnetic equivalent of the acoustic signal) **and sends the audio signal to a remote processor via a communication link**. The **remote processor, such as a server, performs speech recognition processing on the audio signal to determine the command to be carried out and performs voice recognition processing...**If a command is identified and the voice is authenticated, the **remote processor sends a command signal to the local device via a communication link to control the state of the local device.**”)

- 2:65-3:3 (“The local device whose state is altered by the command signal received from the remote processor can include the local command processor, which could be the case, for example, where the local device is a two-way radio (e.g., a military radio), a wireless telephone, or other devices capable of accepting a voice or acoustic command.”)
- 3:15-25 (“Examples of acoustic commands that could be issued to control a local device and processed according to the invention are as follows. Where the local device is a two-way radio or the like, a typical command might be “switch to sub-net 3,” which when interpreted and authenticate by the remote processor would ***alter the state of the radio such that the radio, and possibility other networked radios, are tuned to a particular sub-network.*** Where the local device is a television, a typical command might be “***switch to channel 7***” or “***order pay-per-view program 25.***”)
- 3:30-35 (“The system of the present invention has particular applicability in networked systems where ***control of a local device requires interaction or authorization of a remote party or entity in addition to speech and voice processing of a voice command.***”)
- 4:5-14 (“Referring to FIG. 1, a block diagram for a system for controlling a local device using remote speech and voice recognition processing according to an exemplary embodiment of the present invention includes ***a local command processor 110, a local controllable device 120, a remote processor 130,*** a communication link 140 between the local command processor 110 and the remote processor 130, and a ***communication link 150 between the remote processor 130 and the local controllable device 120.***”)
- 4:32-35 (“Local command processor 110 and local controllable device 120 are typically incorporated into a common local device when the ***local device commonly includes a microphone,*** such as a wireless telephone or a two-way radio”)
- 4:44-49 (“***Local command processor 110 includes a microphone or similar mechanism for picking up acoustic signals.***”)
- 4:53-63 (“In ***operation 210, to control the local controllable device 120, the operator or user at the local site issues an audible voice command (which is an acoustic signal) which can be detected by the microphone of the local command processor 110...*** Once an acoustic signal is

- detected, local command processor 110 converts the acoustic signal to an audio signal, such as an electrical or electromagnetic representation of the acoustic signal (operation 220). The audio signal can be an analog signal or can be sampled to form a digital signal.”)
- 4:64-5:15 (“Communication link 140 can employ any one or combination of communication media and protocols to convey the audio signal from the local site to a remote site (operation 230)...The audio signal can be conveyed via any of a variety of communication schemes including, but not limited to, a conventional land-line telephone network, RF links, private or public data networks (e.g., an intranet, wide area network, or the Internet), satellite communications, and combinations thereof. Likewise, the audio signal can be transmitted as an analog signal or as a digital signal using a suitable transmission protocol (e.g., a TCP/IP compliant protocol, packetized data, VOIP, etc.). Depending on the nature of communication link 140, **local command processor 110 includes suitable means for processing the audio signal for transmission over communication link 140**, such as a modem, an RF transmitter and antenna, etc.”)
 - 5:32-45 (“**Remote processor 130 includes at least a speech recognition processor and a voice recognition processor.** Although not shown in FIG. 1, **remote processor 130 also includes a receiver capable of receiving audio signals from communication link 140 and a transmitter for transmitting resulting command signals over communication link 150** (e.g., a modem). Upon reception of an audio signal, **the speech recognition processor attempts to identify a voice command** within the audio signal (operation 240 in FIG. 2). The speech recognition processing involves performing any of a variety of signal processing techniques for identifying spoken words or sets of words within an audio signal. Such techniques and algorithms are well known in the art and not described here in detail.”)
 - 6:30-45 (“In operation 260, if a voice command is successfully identified via the speech recognition processing and if the voice recognition processing authenticates that the voice is that of an authorized operator, then the **remote processor transmits a command signal to local controllable device 120 (located at the local site) over communication link 150 in order to alter a state of the local controllable device 120.** The command signal is preferably a digital signal having a format that is

- recognizable to local controllable device 120. Communication links 140 and 150 can be the same, two-way communication link or can be separate communication links. Communication link 150 can employ any one or combination of communication media and protocols, such as those previously listed for communication link 140; however, in general, communication link 150 need not necessarily be of the same type as link 140.”)
- 6:56-7:2 (“The following are some ***examples of states of local controllable devices that can be altered via a command signal***. It will be appreciated that the invention is not limited to these specific examples. In the context of a two-way radio, such as a military or public safety radio, ***the command signal may instruct the radio to change an operational status (e.g., on/off, standby mode, sleep mode, etc)*** or change a setting (e.g., increase/decrease volume, change transmission settings, change frequency, etc.). If the local device is a television, virtually any of the conventional operations that can be performed via a remote control can be carried out via a remotely processed voice command (e.g., ***changing channels, changing volume, on/off, adjusting setting, recording, adjusting screen settings, blocking channels, etc.***).”)
 - 7:15-36 (“[A] local operator (customer) issues a voice command at the local site to control the local controllable device 120, and the local command processor 110 generates the audio signal in the manner described above (operations 410 and 420 in the flow diagram of FIG. 4). In operation 430, ***the local command processor transmits the audio signal to the service provider 170 at a service provider site via a communication link***. The service provider then forwards the audio signal to the remote processor 130 at the remote site (operation 440), which performs speech recognition processing and voice recognition processing as described above (operations 450 and 460). In operation 470, if a valid voice command is identified in the audio signal and if the operator is authenticated as an authorized operator, remote processor 130 transmits a command signal to the service provider. The service provider then forwards the command signal to the local controllable device 120 at the local site to control the state of the device, as described above (operation 480).”)
 - 7:54-63 (“The ***remote processor, which provides speech and voice recognition processing...***”)

104. Wittenberg discloses a local device 160 that includes a local command processor 110 and local controllable device 120. Wittenberg, 4:26-31. “Local command processor 110 includes a microphone or similar mechanism for picking up acoustic signals.” Wittenberg, 4:44-45. Local command processor 110 transmits audio signals containing control commands to remote processor 130, in which speech recognition processor 150 analyzes the audio signal by “attempt[ing] to identify a voice command within the audio signal.” Wittenberg, 5:38-40. Remote processor 130 then sends corresponding control commands (“control signals”) to controllable device 120, “in order to alter a state of the local controllable device 120.” Wittenberg, 6:30-36. Wittenberg provides some non-limiting examples of “states of local controllable devices that can be altered via a command signal:” “instruct the radio to change an operational status (e.g., on/off, standby mode, sleep mode, etc) or change a setting (e.g., increase/decrease volume, change transmission settings, change frequency, etc.),” or “changing channels, changing volume, on/off, adjusting setting, recording, adjusting screen settings, blocking channels, etc.” Wittenberg, 6:56-7:2. Thus, altering the state of the local controllable device 120, as its name suggests, controls the “controllable” device 120. Thus, a person of ordinary skill in the art would have understood that, by altering the state of controllable device 120, control of controllable device 120 (which is a component of local device 160) is initiated.

105. As I described in §IX.A.3, a person of ordinary skill in the art would have been motivated to apply Wittenberg’s teaching of remote processing and instruction transmission to a local device to Jannard’s audio device.

i. Claim [2]

106. **Claim [2] recites “The wearable device of claim 1, wherein the interaction element and logic circuit are configured to initiate control of access to the media content played back at the wearable device or mobile device coupled thereto.”** See my discussion of the elements of claim 1 above (§IX.A). The added limitation of claim [2] is met by Jannard.

107. Jannard discloses that the interaction element (*see* element [1.f]) and logic circuit (*see* element [1.d]) are configured to initiate control of access to the media content played back at the wearable device or mobile device coupled thereto (*see* element [1.g]; “perform any one or a combination of various functions with respect to an audio data file” including “randomize file selection, load a playlist”). *See, e.g.*, Jannard:

- [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.”)

- [0135] (“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". Suitable electronics for enabling and amplifying MP3 storage and playback are well known in the art, and may be commercially available from Sigmatel, Inc. or Atmel, Inc.”)
- [0332] (“The audio file is stored in the audio device 304 at block 354. At block 356, the method 350 determines whether ***the user has instructed the audio device 304 to play back the audio file.*** If not, the method 350 continues to block 358, where the method 350 determines whether ***the user has instructed the audio device 304 to load an audio file into the audio device 304.***”)
- [0333] (“If at block 356 the user has instructed the audio device 304 to play back an audio file, the method 350 continues to block 360. At block 360 the method processes an audio file. In one embodiment, block 360 includes any one or a combination functions that may be performed with respect to an audio file. For example, ***at block 360 the audio file may be selected, the playback volume may be adjusted, the tone, balance, bass, treble, or other audio parameter may be adjusted, and/or any other processing function may occur.***”)
- Fig. 22

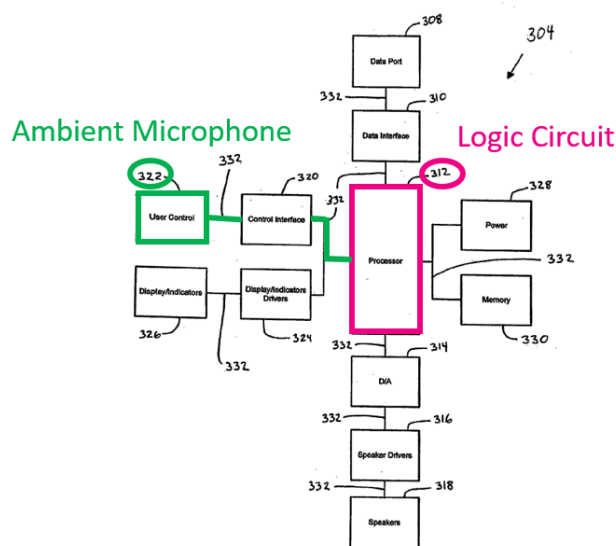


Figure 22

- Fig. 23

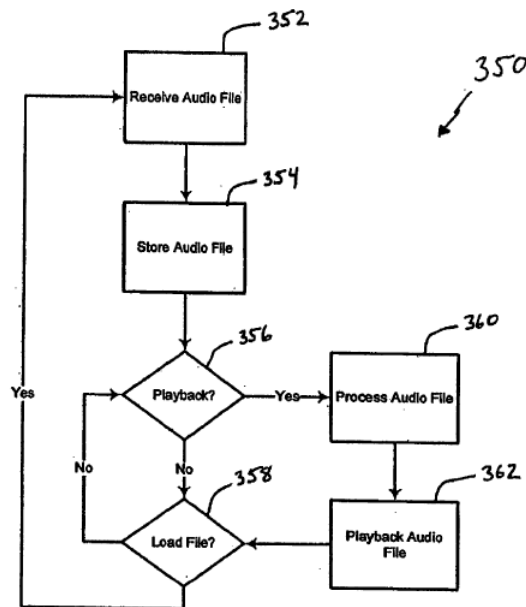


Figure 23

- [0153]: “In the illustrated embodiment, the buttons 73 a and 73 b for controlling the volume of the sound output from the speakers 14A' , 16A' . For example, the **button 73 a can be used for increasing volume and the button 73 b can be used for decreasing volume**. Alternatively, the button 73 b can be for increasing volume and the button 73 a can be for decreasing volume. When a **wearer of the audio device 10A' presses one of the buttons 73 a, 73 b, a simple on-off signal can be transmitted to the device 286. The device 286 can be configured to interpret the on-off signals from the buttons 73 a, 73 b as volume control signals and adjust the volume to the speakers 14A' , 16A' accordingly.**”
- [0154]: “Optionally, a third **command can be generated by pressing both of the buttons 73 a, 73 b simultaneously**. For example, but without limitation, the **device 286 can be configured to interpret simultaneous signals from both the buttons 73 a, 73 b, as a signal for turning on and off an additional feature**. For example, but without limitation, the additional feature can be a **bass boost feature which increases the bass of the audio signal transmitted to the speakers 14A' , 16A' .** Of course, other functions can be associated with the buttons 73 a, 73 b.”

- [0155]: “The **buttons 73 c, 73 d, 73 e can be figured to operate switches to transmit control signals to the device 286 similarly to the buttons 73 a, 73 b.** For example, but without limitation, the button 73 c corresponds to a **power button.** For example, the device 286 can be configured to recognize a signal from the button 73 c as a power on or power off request. In this embodiment, when the device 286 is off, and a signal from the button 73 c is received, the device 286 can turn on. Additionally, the device 286, when in an on state, can be configured to turn off when a signal from the button 73 c is received. Optionally, the device 286 can be configured to, when in an off or standby state, turn on and begin to play an audio file when a signal from the button 73 c is received. Additionally, the device 286 can be configured to pause when another signal from the button 73 c is received. In this embodiment, the device 286 can be configured to turn off only if the button 73 c is held down for a predetermined amount of time. For example, the device 286 can be configured to turn off if the button 73 c is held down for more than two seconds or for three seconds or for other periods of time.”
- [0156]: “The **buttons 73 d and 73 e can correspond to forward and reverse functions.** For example, the button 73 d can correspond to a track skip function. In an illustrative but non-limiting example, such a track skip function can cause the device 286 to skip to a next audio file in the memory of the device 286. Similarly, the button 73 e can correspond to a reverse track skip function in which the device 286 skips to the previous audio file.”
- [0157]: “Optionally, the **buttons 73 d, 73 e can be correlated to fast forward and rewind functions.** For example, the device 286 can be configured to fast forward through an audio file, and play the corresponding sounds at a fast forward speed, when the button 73 d is held down and to stop and play the normal speed when the button 73 d is released. Similarly, the device 286 can be configured to play an audio file backwards at an elevated speed, when the button 73 e is held down, and to resume normal forward play when the button 73 e is released. This arrangement of the buttons 73 a, 73 b, 73 c, 73 d, 73 e provides certain advantages noted above. However, other arrangements of the buttons 73 a, 73 b, 73 c, 73 d, 73 e and the corresponding functions thereof can be modified.”

108. See my discussion for element [1.d], [1.f] and [1.g] above (§IX.A).

109. As I discussed in element [1.d] above, Jannard teaches “processor 312” operates together with “user controls 322.” As I explain above for element [1.f], Jannard also teaches “user controls 322,” including a “button, dial..., microphone, and/or any other input device that may be used by a user to control the audio device.” The control of the device includes “perform[ing] any one or a combination of various functions with respect to an audio data file.” Jannard, [0317]. Jannard offers non-limiting examples of various functions including: “randomize file selection, load a playlist, set a playlist, delete a playlist, repeat playback of all audio files, selected audio files, or a playlist.” Jannard, [0317].

j. Claim [3]

110. **Claim [3] recites “The wearable device of claim 2, wherein the memory storage unit and logic circuit are remote from the housing.”** See my discussion of claim [2] above (§IX.A). The added limitation of claim [3] is met by Jannard.

111. Jannard discloses the memory storage unit (*e.g.*, “memory module 330”; *see* [1.e]) and logic circuit (*e.g.*, “processor 312”; *see* [1.d]) are remote from the housing (*e.g.*, “source electronics such as the MP3 player...may be off board, or located remotely from the audio device”). *See, e.g.*, Jannard:

- [Abstract]: “***Source electronics may be carried by the support or carried remotely***, and in wireless communication with the support using Bluetooth protocols or other data transmission technology.”

- [0012]: “The music source may either be carried by the support, or ***exist remotely from the support and communicate wirelessly with the support.*** In one embodiment, the music source comprises an MP3 player.”
- [0177]: “In general, the component configuration of FIG. 8 enables the audio device 10C to carry interface electronics with the user, such as audio output and audio input. ***However, the source electronics such as the MP3 player, cellular phone, computer or the like may be off board, or located remotely from the audio device 10C.*** This enables the audio device 10C to accomplish complex electronic functions, while retaining a sleek, low weight configuration. Thus, the audio device 10C is in communication with the off board source electronics device B.”)
- [0190]: “In general, the data input management system is configured to allow a 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. ***This allows the wearer to obtain the benefits of multiple input sources, without the need to replace or make any changes to the interface.***”
- [0200]: “***The source electronics may be external to the wearable electronic interface 601, as illustrated in FIG. 8A,*** in which case the communication link 606 may either be a direct electrical coupling (for example hard wiring, or inductive coupling through the body), or wireless.”
- Fig. 8A

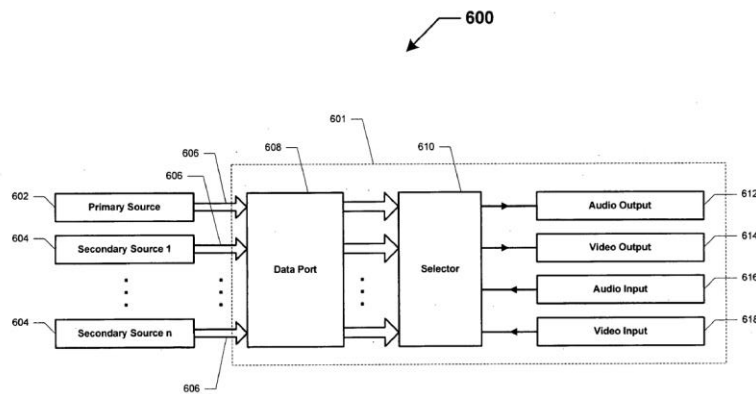


FIG. 8A

- [0322]: “***The memory module 330 may be permanently contained within the audio device 304, or may be removable.*** For example, in one embodiment, the memory module 330 includes an SD memory card, a compact flash memory card, a USB drive, a MEMORYSTICK, SMARTSTICK, and/or any other removable memory device as is well known to those of skill in the art.”
- [0202]: “The source electronics may alternatively be ***carried by or internal to (carried in a cavity or alternatively embedded within) the wearable electronic interface 601.*** For example the primary source 602 may include a digital audio player, such as an mp3 player or other memory device, ***which is attached to or located within the frame of a pair of eyeglasses.*** Electronically-enabled eyewear as a wearable electronic interface 601 is described in greater detail herein. The secondary source may be a cell phone, GPS device or other external device which is in radio communication as needed, with the interface. The primary and secondary sources ***can both be completely contained on the wearable interface, such as built into or carried by a pair of eyewear.***”
- [0204]: “In further embodiments, the source electronics may include a ***video source, or an audio/video source.*** For example, in one embodiment, the source includes a camera for real time viewing of a remote location or viewing direction, or a video playback device such as a DVD or VCR or solid state video storage device. The source electronics may alternatively be a tuner, a television receiver, or any other device capable of providing a signal indicative of still or moving images. In one embodiment, the primary source 602 provides a

photograph, a video clip, an email, a videomail or a voicemail message in accordance with any of the embodiments described herein.”

- [0216]: “The ***selector 610 may be located on the wearable electronic interface 601, or may be remote from it.*** For example, in one embodiment, the wearable electronic interface 601 includes a pair of electronically enabled eyeglasses, and the selector 610 comprises a manually activated control such as a button or touch pad located on an earstem, an orbital or the bridge, or on a ***remote associated component such as the cell phone or a wristwatch. Any other wearable electronic interface 601 or selector 610 location may be utilized.***”

112. See my discussion for element [1.a], and [1.d]-[1.e]. above (§IX.A).

113. Jannard teaches the “memory module 330 may be permanently contained within the audio device 304, or may be removable.” Jannard, [0322]. Jannard further teaches “source electronics such as the MP3 player...may be off board, or located remotely from the audio device 10C,” where the source electronics such as the MP3 player includes a processor for playing music. Jannard, [0177]. Thus, the memory module and processor are remote from the housing of the audio device (*see* [1.a]).

k. Claim [4]

114. **Claim [4] recites “The wearable device of claim 2, wherein the memory storage unit or logic circuit or both are remote from the housing.”** See my analysis of claim [2] above. The added limitation of claim [4] is met by Jannard.

115. As I explain above in connection with claim 3, Jannard teaches that

both the memory module 330 and processor 312 are remote from the housing.
Thus, at least one of the memory module 330 and processor 312 is remote from the housing.

l. Claim [5]

116. **Claim [5] recites “The wearable device of claim 1, wherein the wearable device comprises eye-wear.”** See my discussion of claim [1] above (§IX.A). The added limitation of claim [5] is met by Jannard.

117. Jannard discloses the wearable device comprises eye-wear (e.g., “eyeglass,” see [1.a]).

118. As I discussed in element [1.a], Jannard teaches “an eyeglass,” which is illustrated exemplarily in Figure 12.

m. Claim [6]

119. **Claim [6] recites “The wearable device of claim 1, wherein the interaction element comprises a voice controlled user interface.”** See my discussion of claim [1], particularly elements [1.d] and [1.f], above (§IX.A). The added limitation of claim [6] is met by Jannard.

120. Jannard discloses the interaction element comprises a voice controlled user interface (e.g., “user controls 322” including a “microphone” (see [1.d], [1.f], [1.g])).

121. See my discussion of elements [1.d], [1.f], and [1.g] above (§IX.A).

122. As I discussed in element [1.f], Jannard teaches an interaction element in the form of “user controls 322” that include a “button, dial ..., microphone, and/or any other input device that may be used by a user to control the audio device.” Jannard, [0316].

n. Claim [7]

123. **Claim [7] recites “The wearable device of claim 1, wherein the wearable device further comprises a button with multiple contacts coupled to the logic circuit that forms a portion of the interaction element.”** See my discussion of claim [1], above (§IX.A). The added limitation of claim [7] is met by Jannard.

124. Jannard discloses the wearable device further comprises a button with multiple contacts (*e.g.*, “3-way button 152,” with “[a]ppropriate contacts...mounted beneath the button 152 so as to be activated individually according to the modes of operation” (*see* [1.f])) coupled to the logic circuit (*e.g.*, “processor 312 to receive commands from a user via the user controls 322”) that forms a portion of the interaction element (*e.g.*, “user controls 322 include a button” (*see* [1.f])). *See, e.g.*, Jannard:

- Fig. 11

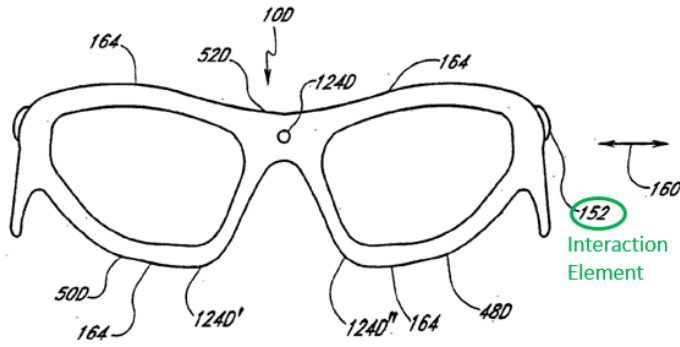


FIG. 11

- Fig. 22

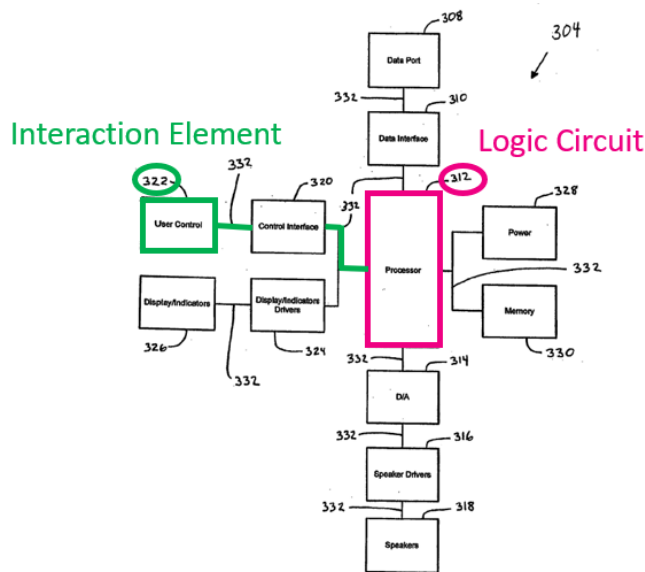


Figure 22

- [314]: “Audio device 304 generally includes...**processor 312**,...**user controls 322**... however, any one or more of these components may be combined. For example, in one embodiment, data interface 310, control interface 320, display/indicator drivers 324, digital-to-analog converter 314, and speaker drivers 316 are combined with processor 312 into a single component.”
- [0257]: “In the illustrated embodiment, the user interface device 150 is in the form of a **3-way button**. **The 3-way button 152 is configured to have three modes of operation...Appropriate contacts can be mounted beneath the button 152 so as to be activated individually according to the modes of operation.**”

- [0258]: “In one illustrative and non-limiting embodiment, ***the button 152 can be used to control volume.*** For example, by pressing on the forward portion 156, a contact can be made causing the transceiver 114 or the interface 116 to increase the volume of the speakers 14D, 16D. Additionally, by pressing on the rearward portion 158 of the button 152, the transceiver 114 or interface 116 could lower the volume of the speakers 14D, 16D.”
- [0123]: “With reference to FIG. 3I, the ***audio device 10A ’ can include at least one button 73 a.*** In the illustrated embodiment, the audio device 10A ’ includes ***five buttons***; a first button 73 a and a second button 73 b mounted to the left ear stem 54 ’ , and a third button 73 c, a fourth button 73 d, and a fifth button 73 e mounted to the right ear stem 56 ’ . Of course, this is one preferred embodiment of the arrangement of the buttons 73 a, 73 b, 73 c, 73 d, 73 e. ***Other numbers of buttons and other arrangements of buttons are also applicable.***”
- [0126]: “Further, by disposing the ***button 73 a on an upper portion of the ear stems 54A ’*** , and by providing the ear stems 54A ’ with an opposite lower surface that faces an opposite direction relative to the upper surface, a wearer can grasp the ear stems 54A ’ from the side, as illustrated in FIG. 38, thereby allowing the user to counteract the actuation force required to actuate the button 73 a without having to insert a finger between a side of the wearer's head and ear stems 54A ’ . In any of the embodiments herein, the surface which opposes the buttons may be provided with any of a variety of tactile feedback structures, such as ridges or bumps, that have a predetermined alignment with respect to the buttons. This can assist the user in positioning their thumb in the identical position each time, so that the user, after a learning period, can rapidly reach for the controls, position their hand and identify with which button their fingers are aligned. See, for example, the tactile indicium illustrated in FIG. 24.”
- [0127]: “FIG. 3J illustrates an exploded view of an exemplary embodiment of the audio device 10A ’ . As shown in FIG. 3J, the left side ear stem 54A ’ defines an electronic housing portion 250 which defines an internal cavity 252 configured to receive electronic components. The electronics housing 250 includes an upper surface 254 and lower surface 260. The upper surface 254 extends generally

outwardly from the ear stems 54A' and around the internal cavity 252. The upper surface also includes apertures 256, 258 through which **buttons 73 a, 73 b**, respectively, extend."

- [0128]: "The housing 250 includes a lower surface 260. The lower surface 260 (which may contain ridges, apertures or slots) faces in an opposite direction from the upper surface 254 of the housing 250. Preferably, the lower surface 260 is at least about 0.25 inches, and may be 0.5 inches or 0.75 inches or more wide. As such, the lower surface 260 provides a surface which allows a wearer to easily grasp the ear stem 54A' so as to balance an actuation force supplied to the **button 73 a, 73 b**."
- [0129]: "A cover member 262 cooperates with the housing 250 to define the closed internal cavity 252. In the illustrated embodiment, the internal cavity 252 includes at least one compartment **configured to receive an electronic circuit board 264 which includes at least one switch for each of the buttons 73 a, 73 b. In an exemplary but non-limiting embodiment, the board 264 can include two switches, one for each of the buttons 73 a, 73 b, which are configured to control a volume output from the speakers 14A' , 16A' .** The cover 262 can be attached to the ear stem 54A' with any type of fastener, such as, for example, but without limitation, screws, rivets, bolts, adhesive, and the like."
- [0214]: "The selector 610 may include a mechanical, electrical, or an electro-mechanical switch. For example, in one embodiment, **the selector 610 includes a manually operable control such as a** toggle switch, a rocker switch, a jumper, a dial, a **button**, a knob, or a combination thereof. In another embodiment, the **selector 610** includes an electronically operable control such as a transistor, a bank of transistors, a relay, a **circuit, logic**, a RAM, a ROM, a PLD, an EPROM, an EEPROM, a **microprocessor**, a microcontroller, a multiplexor, a demultiplexor, or a combination thereof. 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]: "In use, the wearer may be listening to music from the primary source when an incoming phone call is received. If the incoming call is from a number which has been preprogrammed into the system, the

override function allows the name of the caller or other information associated with that incoming phone number to be presented audibly over the speakers. ***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.***

125. See my discussion of elements [1.d], and [1.f] above (§IX.A).

126. Jannard teaches “user controls 322 includ[ing] a button” coupled to “processor 312,” which receives commands from a user via the user controls 322. Jannard, [0315]-[0316]. One of the buttons that form user controls 322 is “3-way button 152,” and “[a]ppropriate contacts can be mounted beneath the button 152 so as to be activated individually according to the modes of operation.” Jannard, [0257]. Jannard offers a non-limiting example of button 152 “used to control volume” “by pressing on the forward portion” and “rearward portion.” Jannard, [0258].

127. Further, to the extent that this limitation requires the logic circuit form a portion of the interaction element, Jannard discloses in one embodiment “control interface 320” is “combined with processor 312.” Jannard, [0304]. Thus, button 152 performs operations with the assistance of processor 312, since button 152 is part of user controls 322, which are in turn coupled to control interface 320. A person of ordinary skill in the art would have understood that the button 152 and its underlying circuitry perform its operations with the assistance of processor 312

because the processor would receive the input of the user via the button 152, would process that input, and then output a corresponding command to the appropriate hardware output to control that component. As seen in Figure 22, user controls 322 are connected to the processor 312, and as such, the processor receives the input of the user controls, including button 152. Thus, processor 312 would be considered part of the interaction element. At the very least, it would have been obvious to a person of ordinary skill in the art to group the processor and button to cooperatively control the modes of operation, as shown in Figure 22. Such a grouping would reduce bulk, as compared to, e.g., a device with a separate processor dedicated to button 152 to process the inputs.

(1) **Claim [8]**

128. **Claim [8] recites “The wearable device of claim 1, wherein the initiating of the operations of the wearable device is in response to a trigger of a predetermined alert condition.”** This element is met by Jannard. See my discussion of claim [1], above (§IX.A). The added limitation of claim [8] is met by Jannard.

129. Jannard discloses the initiating of the operations of the wearable device (*e.g.*, “initiate a telephone call”) is in response to a trigger of a predetermined alert condition (*e.g.*, “when a caller, such as a supervisor, a customer, or other priority individual attempts to contact the wearer”). *See, e.g.*,

Jannard:

- [0237] (“Whether or not the wearer has the ability to prevent a priority, secondary source 604 from interrupting the primary source 602 may be configured based upon the user environment. For example, in a commercial or industrial environment, the primary source 602 may include an mp3 player, and the secondary source 604 may include a paging system, or telephone intercom system. In such embodiment, ***a wearer of a wearable electronic interface 601, such as a production-line worker, security guard, office employee, etc., may use the wearable electronic interface 601 to listen to mp3 audio or other entertainment. However, when a caller, such as a supervisor, a customer, or other priority individual attempts to contact the wearer via the secondary source 604, the wearable electronic interface 601 automatically accepts the incoming communication request***, and allows the caller to communicate with the wearer as discussed herein.”)

130. See my discussion of element [1.f] above (§IX.A).

131. Jannard teaches that a user “configure[s]” the audio device such that, “when a caller, such as a supervisor, a customer, or other priority individual attempts to contact the wearer,” the audio device “automatically accepts the incoming communication request, and allows the caller to communicate with the wearer.” Jannard, [0237]. The receipt of a call from a high priority individual such as a supervisor or customer is an alert condition predetermined by the device. For example, the device may be “configured based upon the user environment” and preferences. Jannard, [0237]. To set this configuration, a person of ordinary skill in the art would have understood that a user uses “user controls 322” (the interaction element), as this is the mechanism by which a user interacts with the audio device, as the name suggests.

o. Claim [9]

132. **Claim [9] recites “The wearable device of claim 1, wherein the wearable device comprises an earphone.”** See my discussion of claim [1], above (§IX.A). The added limitation of claim [9] is met by Jannard.

133. Jannard discloses the wearable device comprises an earphone (*e.g.*, “speakers 14D, 16D”). *See, e.g.*, Jannard:

- Fig. 12 (annotated)

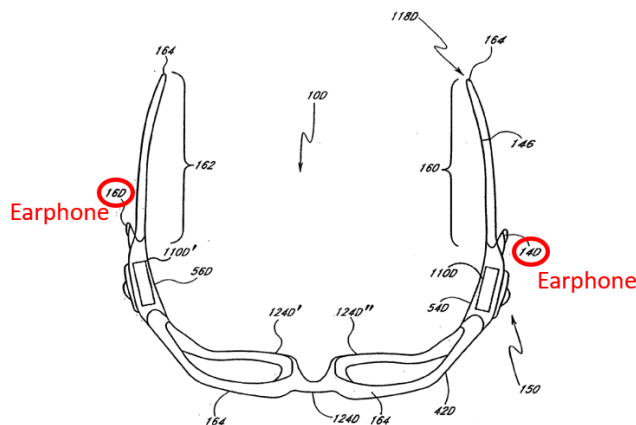


FIG. 12

- [0258] (“[B]y pressing on the forward portion 156, a contact can be made causing the transceiver 114 or the interface 116 to increase the volume of the *speakers 14D, 16D.*”)
- [0262] (“In another alternative, the power contacts 164 are disposed on the upper edges of the orbitals 48D, 50D. In this configuration, the audio device 10D is laid in a cradle in an inverted position, such that the contacts 164 make electrical contact with corresponding contacts in the cradle. This position is advantageous because it prevents weight from being applied to the supports 28D, 30D. This prevents misalignment of the speakers 14D, 16D.”)

134. See my discussion of element [1.pre], above (§IX.A).

135. Jannard teaches audio device 10D includes “speakers 14D, 16D.”

Jannard, [0258]. Further, speakers 14D, 16D are earphones because they are located over the ear. Jannard, [0262]. A person of ordinary skill in the art would understand that these are earphones from their described functionality. For example, when Jannard teaches the placement of certain circuitry “prevents misalignment of the speakers 14D, 16D,” the “misalignment” referred to is with the “ears,” and is intended to ensure the speakers are located directly over the ears to play sound towards the ears of the user. Jannard, [0262]. Thus, the speakers are earphones. Further, to the extent further disclosure of the speakers being directly over the ears is required, it would have been obvious to extend the speakers to be placed directly over the ear to minimize the escape of sound and ensure others do not hear the audio being played for the user. For example, by placing the speaker directly over the ear, the volume can be minimized, such that the sound from the speaker is less likely to travel far enough from the ears to be heard by others.

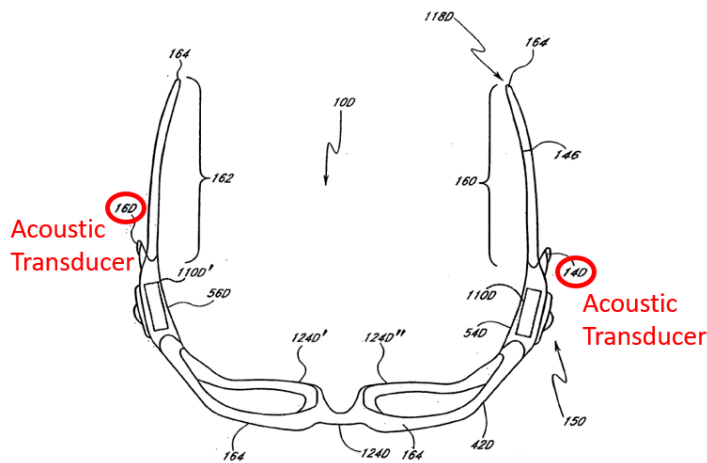
p. Claim [10]

136. **Claim [10] recites “The wearable device of claim 1, wherein the wearable device comprises at least a portion of a personal computer system, a personal music player system, a personal monitoring system, an automotive audio system, a home audio system, an avionics audio system, a personal video system, a mobile cell phone system, a personal digital assistant system, a standalone accessory, or an advanced eye-wear system with acoustical**

transducers.” See my discussion of claim [1], above (§IX.A). The added limitation of claim [10] is met by Jannard.

137. Jannard discloses the wearable device comprises at least a portion of...an advanced eye-wear system with acoustical transducers (e.g., “eyeglass” with “speakers 14D, 16D”). See, e.g., Jannard:

- Fig. 12



- [0255]: “[T]he stem 56D would also become part of the antenna 118D, 118D’ , thereby allowing transmission and reception of signals on three sides of the user’s head. Thus, where at least a portion of a frame of an **eyeglass** is used as the antenna for the wireless transceiver 114, the audio device benefits from enhanced antenna efficiency.”
- [0258]: “[I]ncrease the volume of the **speakers 14D, 16D.**”
- [0176]: “[T]he **user U can listen** to the audio output of the cellular phone, such as the voice of a caller, **through sound transducers....**”
- [0134]: “Additionally, the printed circuit board 284 includes **an audio file storage and playback device 286.**”
- [0135]: “The device 286 can be configured to store and playback any desired type of electronic audio and/or video file....For example, **the**

audio storage and playback system can be configured to store MP3 files in a memory and to play back the MP3 files through the speakers 14A', 16A'.

- [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. ***The interface may comprise any of a variety of physical forms, including electronically enabled eyeglasses.*** 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.”
- [0012]: “The wearable cell phone and music system may additionally comprise radiofrequency electronics such as a Bluetooth receiver or transceiver, for wireless communication between the ***cell phone*** and the support. The system may additionally comprise radiofrequency electronics for wireless communication between the music source and the support. The music source may either be carried by the support, or exist remotely from the support and communicate wirelessly with the support. In one embodiment, the music source comprises an MP3 player.”

138. See my discussion of elements [1.pre] to [1.e], above (§IX.A).

139. Jannard teaches audio device 10D is an “eyeglass” with speakers 14D, 16D,” as I discussed in element [1.pre]. Speakers are a form of acoustical transducers. ’692, 5:14-18 (identifying “microphones” as an example of acoustical transducers); Ex. 1018 (Spoken language Processing, Huang, 2001 (“SLP”)), page 482 (“Acoustical transducers are devices that convert the acoustic energy of sound into electrical energy (microphones) and vice versa (loudspeakers).”).

140. Further, Jannard discloses that audio device 10A' “includes an audio file storage and playback device 286,” “configured to store and playback any desired type of electronic audio and/or video file,” where the audio files include “MP3 files.” Jannard, [0134]-[0135]. Thus, the playback device 286 constitutes a “personal music player system.”

141. Further, Jannard describes combined capabilities and functionalities as advantageous functionality in its audio device. Jannard, [0009]-[0010]. Thus, a person of ordinary skill in the art would have understood that Jannard's teachings of integrating playback device 286 apply to audio device 10D. At the very least, a person of ordinary skill in the art would have found it obvious to advantageously provide extra functionality in the audio device 10D. Further, this integration avoids the need to carry an extra device to play music.

q. Claim [11]

142. **Claim [11] recites “The wearable device of claim 1, wherein the wearable device further comprises speaker for providing an audible output.”** See my discussion of claim [1], above (§IX.A). The added limitation of claim [11] is met by Jannard.

143. Jannard discloses the wearable device further comprises speaker for providing an audible output (*e.g.*, “volume of the speakers 14D, 16D”). *See, e.g.*, Jannard:

- Fig. 12

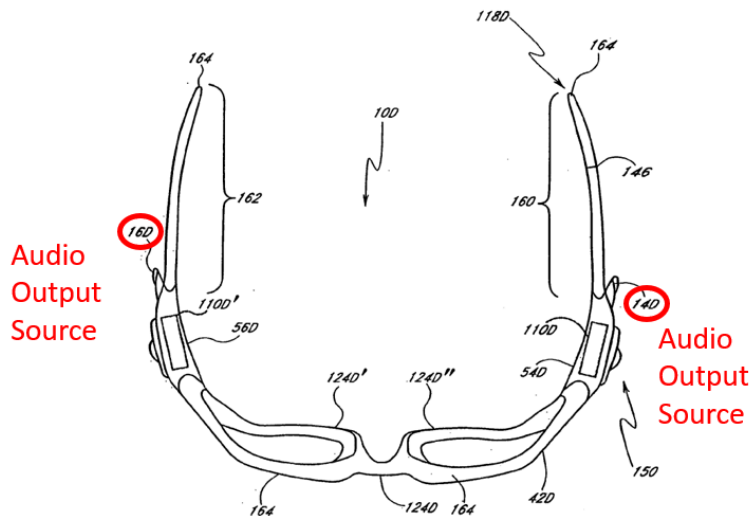


FIG. 12

- [0258]: “In one illustrative and non-limiting embodiment, the button 152 can be used to control volume. For example, by pressing on the forward portion 156, a contact can be made causing the transceiver 114 or the interface 116 to **increase the volume of the speakers 14D, 16D**. Additionally, by pressing on the rearward portion 158 of the button 152, the transceiver 114 or interface 116 could **lower the volume of the speakers 14D, 16D**.”
- [0074]: “The support 12 is configured to be supported by the head 18. Thus, the support 12 can be in the form of any of a variety of wearable garments or devices, such as **any known headwear**. For example, but without limitation, the support 12 can be in the form of a hat, sweatband, tiara, helmet, headphones, visor, and any of a variety of eyewear such as goggles, masks, face shields and eyeglasses.”

144. See my discussion of claim [9], above (§IX.A).

145. Jannard teaches the “volume” of “speakers 14D, 16D” can be adjusted. Jannard, [0258]. Here, the “volume” of the speakers refers to the volume of the audio output from the speakers. Thus, the speakers 14D, 16D have an audible output that is provided to the user.

r. **Claim [13]**

146. **Claim [13]** recites “The wearable device of claim 1, wherein the logic circuit is configured to perform a search of the web in response to a voice command and provide a web search result to the wearable device.” See my discussion of claim [1], above (§IX.A). The added limitation of claim [13] is met by Jannard.

147. Jannard discloses the logic circuit (*e.g.*, “processor;” see [1.d]) is configured to perform a search of the web (*e.g.*, “internet browsing”) in response to a voice command (*e.g.*, “recognize speech from the user, convert the speech to AT commands”) and provide a web search result to the wearable device (*e.g.*, “control this source device,” including “internet browsing”). See, *e.g.*, Jannard:

- [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, 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. 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.”

- [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.”
- [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, 16 E, 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.”

148. Jannard teaches “internet browsing.” Jannard, [0298]. Jannard describes that such internet browsing is done 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]. Accordingly, the audio device 10D would receive the search results of any internet search.

s. Claim [14]

149. **Claim [14] recites “The wearable device of claim 1, wherein the logic circuit is configured to cooperatively initiate control of the media content of the wearable device and initiate operations of the wearable device in response to analysis of the audio captured at the wearable device at the ambient microphone and analyzed at the remote server in response to instructions received from the remote server.”** See my discussion of claim [1], above (§IX.A). The added limitation of claim [14] is met by Jannard.

150. Jannard discloses the logic circuit is configured to cooperatively initiate control of the media content of the wearable device (*e.g., see* [1.g]) and

initiate operations of the wearable device (*e.g.*, “answer an incoming phone call”; initiating a call by saying “call alpha”) in response to analysis of the audio captured at the wearable device at the ambient microphone and analyzed at the remote server in response to instructions received from the remote server (*e.g.*, see [1.g]). See, *e.g.*, Jannard:

- [0236]: “In use, the wearer may be listening to music from the primary source when an incoming phone call is received. If the incoming call is from a number which has been preprogrammed into the system, the override function allows the name of the caller or other information associated with that incoming phone number to be presented audibly over the speakers. ***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.”
- [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.”
- [0187]: “The microphone 124 is connected to the interface 116 through a microphone line 130. Thus, the transceiver 114 can receive audio signals from the microphone 124 through the interface 116. As such, the 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. ***The microphone 124 can also be in any of the forms discussed above with reference to the microphone 75.***”

- [0259]: “In a further illustrative and non-limiting example, the medial-lateral movement of the button 152, along the directions identified by the arrow 160, ***can be used to choose different functions performed by the transceiver 114 or the interface 116.*** For example, an inward movement of the button 152 ***could be used to answer an incoming phone call where the audio device 10D is used as an audio interface for a cellular phone.***”
- [0292]: “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.***”

- [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.”

151. See my discussion of element [1.g], above (§IX.A).

152. The limitation added by claim 14 is identical to element [1.g], except that it further requires “initiate operation of the wearable device.” Jannard teaches the “wearer can then elect whether to accept the call, such as by...voice command,” and thus initiates acceptance of a voice call of the wearable device in response to the voice command. Jannard, [0236]. Thus, for example, by the wearer electing to accept the call by corresponding voice command, the call is initiated at the wearable device.

153. Jannard provides other non-limiting examples of initiating operations of the audio device. For example, “if a user chooses a voice command option, the user could issue commands such as...‘call alpha’.” Jannard, [0297]. Thus, if the user inputs voice command “call alpha,” this would result in the initiation of a phone call operation at the audio device. Jannard, [0297].

t. Claim [16]

154. **Claim [16] recites “The wearable device of claim 1, wherein the**

logic circuit is configured to retrieve specific media content in response to voice commands captured at the wearable device using the ambient microphone, and the wearable device receives results from the remote server in response to speech recognition processing of the voice command at the remote server.” See my discussion of claim [1], above (§IX.A). The added limitation of claim [16] is met by Jannard in view of Wittenberg.

155. Jannard discloses the logic circuit (*e.g.*, “processor,” *see* [1.d]), is configured to retrieve specific media content (*e.g.*, “downloading of music from the content source 302 (*e.g.*, a user's computer) to the audio device 304,” *see* [1.d], [1.f]) in response to voice commands captured at the wearable device using the ambient microphone (*e.g.*, “commands from a user via the user controls 322,” which “include a microphone that receives a voice command”). *See, e.g.*, Jannard:

- [0191]: “[*T*]he source electronics may include a computing device, such as a computer, a **server**, a network, drive, RAM, ROM or other non-removable or removable memory chip.”
- [0303]: “In one embodiment, the audio network 300 is configured for the **downloading of music from the content source 302** (*e.g.*, a user's computer) **to the audio device 304.**”
- [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.”
- [0316]: “In one embodiment, the user controls 322 include a microphone that receives a voice command.”

- [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.”
- [0187]: “The microphone 124 is connected to the interface 116 through a microphone line 130. Thus, the transceiver 114 can receive audio signals from the microphone 124 through the interface 116. ***As such, the 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.*** The microphone 124 can also be in any of the forms discussed above with reference to the microphone 75.”
- [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.***”

156. See my discussion of elements [1.d], [1.f], [1.g] and claim [14].

above (§IX.A).

157. As I discussed in connection with elements [1.d] and [1.f], Jannard teaches “downloading of music from the content source 302...to the audio device 304.” Jannard, [0303]. And as I discussed in connection with element [1.d], Jannard teaches “processor 312...receive[s] commands from a user via the user controls 322”—including “a microphone that receives a voice command”—to “perform any one or a combination of various functions with respect to an audio data file.” Jannard, [0315]-[0317]. Thus, the processor downloads music from a content source in response to commands from the user via user controls 322, which

include a microphone to receive voice commands.

158. Wittenberg discloses a device receives results from the remote server (e.g., “local device 160” receives “a command signal” from “remote processor 130,” see [1.g]) in response to speech recognition processing of the voice command at the remote server (e.g., “remote processor 130” includes a “a speech recognition processor [that] attempts to identify a voice command with the audio signal,” see [1.g]). See, e.g., Wittenberg:

- 5:33-43 (“Although not shown in FIG. 1, remote processor 130 also includes a receiver capable of receiving audio signals from communication link 140 and a transmitter for transmitting resulting command signals over communication link 150 (e.g., a modem). Upon reception of an audio signal, the speech recognition **processor attempts to identify a voice command within the audio signal** (operation 240 in FIG. 2). The speech recognition processing involves performing any of a variety of signal processing techniques for identifying spoken words or sets of words within an audio signal.”)
- 6:30-45 (“In operation 260, if a voice command is successfully identified via the speech recognition processing...then the **remote processor transmits a command signal to local controllable device 120 (located at the local site) over communication link 150 in order to alter a state of the local controllable device 120**. The command signal is preferably a digital signal having a format that is recognizable to local controllable device 120. Communication links 140 and 150 can be the same, two-way communication link or can be separate communication links. Communication link 150 can employ any one or combination of communication media and protocols, such as those previously listed for communication link 140; however, in general, communication link 150 need not necessarily be of the same type as link 140.”)

159. See my discussion of element [1.g], above (§IX.A).

160. Wittenberg discloses that the “remote processor transmits a command

signal,” which is produced by the remote processor 130 after the “speech recognition processor attempts to identify a voice command with the audio signal.” Wittenberg, 5:33-43, 6:30-45. As described in § IX.A.3, a person of ordinary skill in the art would have been motivated to apply Wittenberg’s teachings of remote processing and instruction transmission to a local device to Jannard’s audio device.

u. Claim [17]

161. **Claim [17] recites “The wearable device of claim 1, wherein the wearable device further comprises a presentation device operatively coupled to the wearable device for presenting one or more of text, images, webpages, location information, scheduling information, or audible outputs to complement detected voice commands at the wearable device and speech recognition results received at the wearable device from the remote server in response to analysis of the detected voice commands at the remote server.”**

See my discussion of claim [1], above (§IX.A). The added limitation of claim [17] is met by Jannard in view of Wittenberg.

162. Jannard discloses the wearable device further comprises a presentation device (*e.g.*, “audio output”) operatively coupled to the wearable device for presenting one or more of text, images, webpages, location information, scheduling information, or audible outputs (*e.g.*, “audio output that is audible to the user”) to complement detected voice commands at the wearable device (to provide “audio

indicators back to the user” when “microphone...receives a voice command”).

See, e.g., Jannard:

- [0100]: “The **audio device 10A can be used as an audio output device for any type of device which provides an audio output signal.**”
- [0278]: “The digital-to-audio converter 204 can be connected to the speakers 14,16 and thereby **provide an audio output that is audible to the user.**”
- [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.”
- [0316]: “In one embodiment, the **user controls 322 include a microphone that receives a voice command.**”
- [0015]: “In accordance with a further aspect of the present invention, there is provided a wearable electronically enabled interface system, for providing **audio and or video input to the wearer** from at least two sources. The system comprises a wearable support, for carrying at least a portion of the interface. **At least one data port is carried by the support, for receiving data from at least a first** and optionally a second source. A selector is provided, enabling the wearer to direct data from a desired one of the first and second sources to the interface.”

- [0083]: “In the illustrated embodiment of the audio device 10A, the support 12A is in the form of an eyeglass 40. The eyeglass 40 comprises a frame 42 which supports left and right lenses 44, 46. Although the present audio device 10A will be described with reference to a dual lens eyeglass, it is to be understood that the methods and principles discussed herein are readily applicable to the production of frames for unitary lens eyeglass systems and protective goggle systems as well. Further, the lenses 44, 46 can be completely omitted. Optionally, at least one of the lenses 44, 46 can be in the form of a ***view finder or a video display unit configured to be viewable by a wearer of the support 12A.***”
- [0204]: “In further embodiments, the source electronics may include a ***video source, or an audio/video source.*** For example, in one embodiment, the source includes a ***camera*** for real time viewing of a remote location or viewing direction, or a ***video playback device*** such as a DVD or VCR or solid state video storage device. The source electronics may alternatively be a tuner, a television receiver, or any other device capable of providing a signal indicative of still or moving images. In one embodiment, the primary source 602 provides a ***photograph, a video clip, an email, a videomail or a voicemail message*** in accordance with any of the embodiments described herein.”
- [0225]: “In addition, the selector 610 includes sufficient logic to know that when an incoming telephone call is being received from a source 602, 604, ***the audio input 616 (e.g., microphone) of the wearable electronic interface 601 will be activated to provide communication from the wearer back to the secondary source 604.*** Similarly, if the source electronics selected by the user carries video signals, the selector 610 additionally activates the video display carried by the eyeglasses or other support structure. If the source electronics selected by the user or automatically by the selector 610 includes only an audio signal, the microphone and video display, if present, remain dormant.”
- [0263]: “With reference to FIGS. 8, 9A, and 9B, in another embodiment, the audio device 10C is advantageously adapted to support any of a variety of portable electronic circuitry or devices which have previously been difficult to incorporate into conventional headsets due to bulk, weight or other considerations. For example, but without limitation, the electronics are digital or other storage devices and retrieval circuitry such as for ***retrieving music or other information from MP3 format memory***

or other memory devices. The audio device 10C can carry any of a variety of receivers and/or transmitters, such as transceiver 114. For example, but without limitation, the audio device 10C can carry receivers and/or transmitters for music or for global positioning. In another example, the audio device 10C can carry receivers and/or transmitters for telecommunications (e.g., telecommunications devices). As used herein, the term **“telecommunications devices” is intended to include telephone components as well as devices for communicating with a telephone.** **For example, “telecommunications devices” can include one or more transceivers for transmitting an audio signal to a cellular phone to be transmitted by the cellular phone as the speaker's voice, and/or for receiving an audio signal from a cellular phone representing a caller's voice.** Of course, other audio, video, or data signals can be transmitted between the audio device 10 C and such a cellular phone through such transceivers.”

- [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.”

163. See my discussion of elements [1.f] through [1.g], and claims [11]-[12], and [16], above (§IX.A).

164. Jannard further teaches that the audio device “can be used as an audio output device” and thus “provide[s] an audio output that is audible to the user.”

Jannard, [0100], [0278]. When the “microphone...receives a voice command” ([0316]) it is provided to the “source device,” such as a “server,” which includes a “speech recognition engine...that is properly trained to recognize the voice of the user.” Jannard, [0297]. Moreover, when “the user speaks,” the audio device is “configured to issue audio indicators back to the user...to guide the user through such a voice command.” Jannard, [0296].

165. Wittenberg discloses speech recognition results received at the wearable device from the remote server (“local device 160” receives “a command signal” from “remote processor 130,” *see* [1.g], [12]) in response to analysis of the detected voice commands at the remote server (*e.g.*, “remote processor 130” includes a “a speech recognition processor [that] attempts to identify a voice command with the audio signal,” *see* [16]).

166. See my discussion of elements [1.g] and claim [16], above (§IX.A).

167. As I discussed in [1.g], Wittenberg discloses that local device 160, including local command processor 110, transmits audio signals containing voice commands to remote processor 130, where speech recognition processor analyzes the audio signal. Remote processor 130 then sends the resulting control command to controllable device 120 to initiate various operations. Wittenberg, 4:6-35, 6:30-45.

168. And as I described in §IX.A.3, a POSITA would have been motivated

to apply Wittenberg’s teachings of remote processing and instruction transmission to a local device to Jannard’s audio device.

v. Elements [18.pre]-[18.b]

169. **Element [18.pre]** recites “**A wearable device, comprising.**” This element is met by Jannard. See my discussion of the element [1.pre] above, which is the same limitation (§IX.A).

170. **Element [18.a]** recites “**a housing for the wearable device.**” This element is met by Jannard. See my discussion of the element [1.a] above, which is the same limitation (§IX.A).

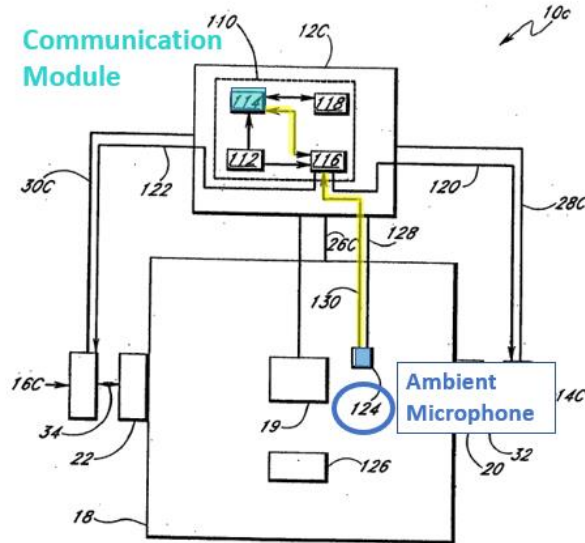
171. **Element [18.b]** recites “**an ambient microphone mounted on or within the housing.**” This element is met by Jannard. See my discussion of the element [1.b] above, which is the same limitation (§IX.A).

w. Element [18.c]

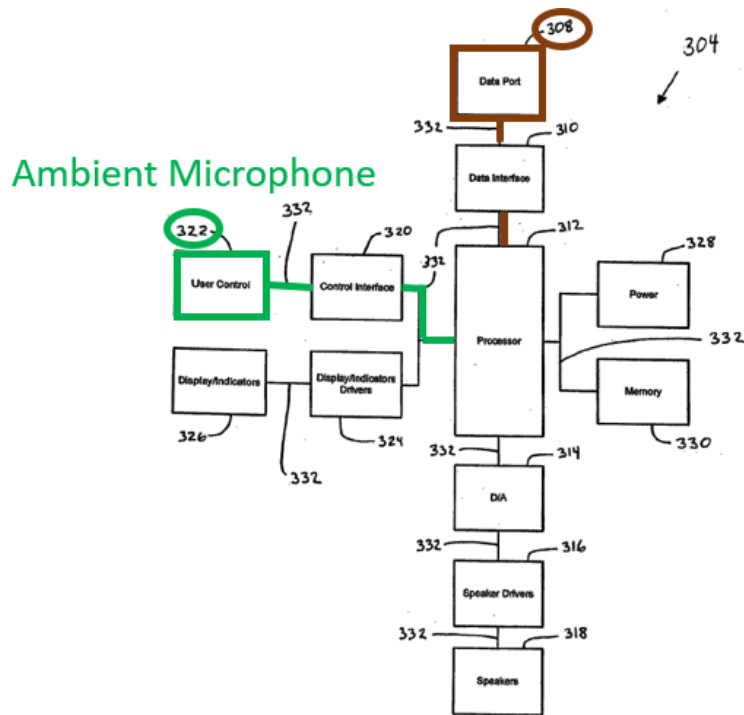
172. **Element [18.c]** recites “**a communication module communicatively coupled to the ambient microphone.**” This element is met by Jannard.

173. Jannard discloses a communication module (*e.g.*, “transceiver 114,” *see* [1.c]) communicatively coupled to the ambient microphone (*e.g.*, “transceiver 114 can be configured to operate in a full duplex mode in which...audio signals from the microphone...124D...are transmitted...to a cooperating wireless device”). *See, e.g.*, Jannard:

- [0288] (“[T]he **transceiver 114 can only...transmit signals, for example, from the microphone 75, 124, 124D, 124E to another device** through the antenna 118, 118D, 118D’ . Alternatively the transceiver 114 can be configured to operate in a full duplex mode in which simultaneous of audio signals are received and transmitted to the speakers 14, 14A, 14B, 14C, 14D, 14E, 16, 16A, 16B, 16C, 16D, 16E and simultaneously **audio signals from the microphone 75, 124, 124D, 124E are transmitted through the antenna 118, 118D, 118D’ to a cooperating wireless device.**”)
- [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....”)
- [0187] (“The microphone 124 is connected to the interface 116 through a microphone line 130. Thus, the **transceiver 114 can receive audio signals from the microphone 124** through the interface 116.”)
- [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**”)
- Fig. 9A (annotated)



- Fig. 22 (annotated)



174. See my discussion for elements [1.c]-[1.d] and [1.g], above (§IX.A).

175. Jannard teaches “transceiver 114...configured to operate in a full

duplex mode in which...audio signals from the microphone 75, 124, 124D, 124E are transmitted through the antenna 118, 118D, 118D' to a cooperating wireless device.” Jannard, [0288]. Thus, transceiver 114 is communicatively coupled to microphone 124D.

176. Further, when Jannard’s audio device is implemented with a microphone bank, as Jannard describes in [0265], a person of ordinary skill in the art would have understood, and at minimum found it obvious, that transceiver 114 is coupled to Jannard’s “banks of microphones.” Jannard [0265]. For example, the bank of microphones would be coupled to the transceiver when the device is processing voice commands to be sent to the server. See my discussion on elements [1.c]-[1.d], and [1.g]. In addition, Jannard teaches that “data port 308” is for “wireless” communication over “communication link,” and illustrates data port 308 is communicatively coupled with user controls 322, which as I discussed in connection with element [1.d] (above), includes the ambient microphones. See also my discussion of element [1.f], above.

x. Element [18.d]

177. **Element [18.d] recites “one or more processors communicatively coupled to the ambient microphone.”** This element is met by Jannard. See my discussion of element [1.d] above, where the “logic circuit” corresponds to the “one or more processors” (§IX.A).

y. Element [18.e]

178. Element [18.e] recites “a computer readable medium communicatively coupled to the one or more processors and containing instructions which when executed by the one or more processors, cause the one or more processors to perform functions comprising initiating control of audio content or initiating at least one operation of the wearable device.” This element is met by Jannard.

179. Jannard discloses a computer readable medium (*e.g.*, “memory module 330”) communicatively coupled to the one or more processors (*e.g.*, “processor 312,” *see* Fig. 22) and containing instructions which when executed by the one or more processors (*e.g.*, “configured to operate,” “providing added functionality”), cause the one or more processors to perform functions comprising initiating control of audio content (*e.g.*, “processor 312 receive[s] commands from a user via the user controls 322” to “perform any one or a combination of various functions with respect to an audio data file”) or initiating at least one operation of the wearable device (*see* [1.f]). *See, e.g.*, Jannard:

- Fig. 22

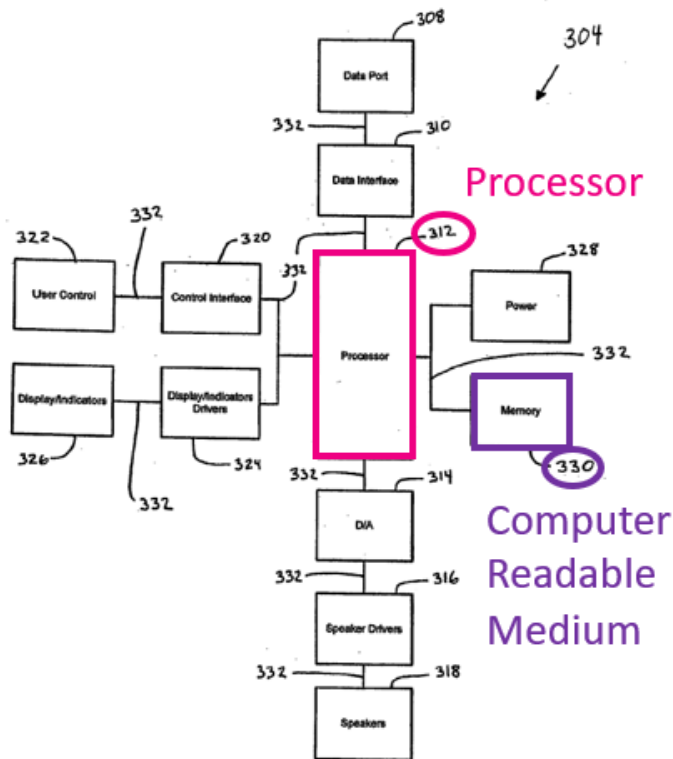


Figure 22

- [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.***”)
- [0289] (“Further, ***the interface 116 can include a processor and a memory for providing added functionality.***”)
- [0307] (“***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).”)

180. See my discussion for elements [1.d]-[1.f] and claims [6] and [10].

above (§IX.A).

181. Jannard teaches the “memory...and the processor are configured to operate together.” Jannard [0135]. Jannard describes the processor and memory operate to “provid[e] added functionality.” Jannard, [0289]. Jannard offers a non-limiting example of the “processor 312...receive[ing] commands from a user via the user controls 322” and, in response, “perform[ing] any one or a combination of various functions with respect to an audio data file.” Jannard, [0315], [0317].

182. Moreover, a person of ordinary skill in the art would have understood that memory module 330 stores instructions causing processor 312 to perform such operations in response to user commands, or that, at the very least, it would have been obvious to include instructions in memory coupled to processor 312 for the reasons I discuss above in connection with element [1.e]. See § IX.A.

z. Element [18.f]

183. **Element [18.f] recites “wherein the one or more processors is configured to cooperatively initiate control of media content of the wearable device in response to analysis of audio captured at the wearable device at the ambient microphone and analyzed at a remote server in response to instructions received from the remote server.”** This element is met by Jannard, in view of Wittenberg. See my discussion of the element [1.g] above, where the “logic circuit” corresponds to the “one or more processors.” (§IX.A).

aa. Claim [19]

184. **Claim [19] recites “The wearable device of claim 18, wherein the one or more processors initiates control of the media content in response to a voice command received and processed by a voice controlled user interface, wherein a remote server receives and analyzes the processed voice command and sends results back to the wearable device.”** See my discussion of the elements of claim [18] above (§IX.A). The added limitation of claim [19] is met by Jannard.

185. As I explain in connection with element [1.g] and claim [6], Jannard discloses the one or more processors (corresponding to “logic circuit”) initiates control of the media content in response to a voice command received and processed by a voice controlled user interface.

186. As I explain in connection with claims [16] and [17], Wittenberg discloses wherein a remote server receives and analyzes the processed voice command and sends results back to the wearable device.

bb. Elements [20.pre]-[20.c]

187. **Element [20.pre] recites “A wearable device, comprising.”** This element is met by Jannard. See my discussion of the element [1.pre] above (§IX.A).

188. **Element [20.a] recites “a wearable structure having at least an ambient microphone.”** This element is met by Jannard. See my discussion of the

elements [1.a] and [1.b] above, where the “housing” is the “wearable structure” (§IX.A).

189. **Element [20.b] recites “a communication module communicatively coupled to the ambient microphone.”** This element is met by Jannard. See my discussion of the element [18.c] above, which is the same limitation (§IX.A).

190. **Element [20.c] recites “one or more processors communicatively coupled to the ambient microphone.”** This element is met by Jannard. See my discussion of the element [1.d] above, where the “logic circuit” corresponds to the “one or more processors” (§IX.A).

cc. Element [20.d]

191. **Element [20.d] recites “a computer readable medium communicatively coupled to one or more processors and containing instructions which when executed by the one or more processors, cause the one or more processors to perform functions comprising initiating control of audio content accessed at the wearable device in response to detected voice commands in the captured audio.”** This element is met by Jannard. See my discussion of element [18.e] and claim 16 above (§IX.A).

192. Jannard discloses a computer readable medium communicatively coupled to one or more processors and containing instructions which when

executed by the one or more processors, cause the one or more processors to perform functions comprising initiating control of audio content (*see* [18.e]) accessed at the wearable device in response to detected voice commands in the captured audio (*see* [16]).

dd. Element [20.e]

193. **Element [20.e] recites “wherein the one or more processors is configured to cooperatively initiate control of media content of the wearable device in response to analysis of audio captured at the wearable device at the ambient microphone and analyzed at a remote server in response to instructions received from the remote server.”** This element is met by Jannard. See my discussion of the element [18.f] above, which is the same limitation (§IX.A)

B. GROUND 2: GROUND 1 IN FURTHER VIEW OF RELAN (Claim 12):

1. Overview of Relan

194. Relan describes a device capable of “exchanging media” information between electronic devices. Relan, Abstract. Relan provides that such information exchanged includes that of global positioning systems through a peer-to-peer network. Relan, Abstract, [0059], [0085]. Relan’s device is applicable to a “mobile handset” with a GPS receiver. Relan, [0295]; [0085].

195. Relan’s device includes a “media exchange network 100” comprising

a “home/location network 107.” Relan, [0048]. The home/location network 107 includes the necessary components for “providing peer-to-peer networking capability.” Relan, [0059]. A “peer-to-peer network,” as Relan explains, “refer[s] to a communication network in which information is transferred or shared by one user with another, using the same program or type of program (i.e., peers).” Relan, [0059]. Relan teaches using “a global positioning system (GPS) receiver” to determine location, and using that location information as “security information.” Relan [0084]. “The position/location information may comprise, for example, coordinate information from a global positioning system (GPS) receiver.” Relan, [0085]. Such information is used when forming a peer-to-peer connection to authenticate a user request. Relan, [0084]-[0085]. “Through the use of such a mechanism, the exchange of multimedia information may be limited to authenticated users, and may further be limited to those users attempting access from pre-determined geographic locations.” Relan, [0085].

2. Motivation to Apply the Relan’s Teachings to Jannard and Wittenberg

196. Jannard, Wittenberg and Relan are analogous art. Jannard, Wittenberg and Relan are in the same field as the ’692 patent, including wearable devices. Jannard, Wittenberg and Relan are reasonably pertinent to alleged problems addressed by the ’692 patent of combining the capacity of several devices into one wearable audio device. E.g., ’692 patent, 1:22-34; Jannard,

Abstract, [0189]; Relan, Abstract, [0042] (disclosing devices that support the “exchange and sharing of multimedia information between, for example, family members and friends in an efficient, user-friendly, and economically viable manner over a closed and secure multimedia exchange network...”); see my discussion of the motivation to combine Jannard and Wittenberg above, §IX.A.3. Relan describes “supporting secure exchange of multimedia information” across devices. Relan, [0025]. This is pertinent to the problem addressed by the ’692 patent because input from multiple data sources, such as audio content playing devices and headphone systems, are not readily compatible. ’692, 1:30-35. Jannard also describes such issues and a desire to solve for compatibility and bulk of multiple devices. See Jannard, [0009]-[0010].

197. A person of ordinary skill in the art would have been motivated to apply Relan’s teachings of using a peer-to-peer network to Jannard’s wearable device (as modified by Wittenberg) to, for example, advantageously allow a user to acquire and share audio files from others. Relan, Abstract, [0171].

198. A person of ordinary skill in the art would have further been motivated to apply Relan’s teachings of using location information to establish a “peer-to-peer” network to Jannard’s wearable audio device because using location information in establishing such a network allows the network to confirm a user is in an appropriate location. Relan, [0054]. This enhances security. As Relan

describes, the physical location of the mobile terminal acts as an “element in the verification of authority” regarding the content the user is requesting, furthering the security of the information. Relan, [0054]; [0084]. Additionally, the added functionality of mobile location authentication allows a user to move around and still access desired information when at an acceptable location, rather than being stuck at a fixed location. *See, e.g.*, Relan, [0020].

199. Additionally, Jannard already teaches a “global positioning system” and it would have been straightforward for the audio device 10D to connect to a peer-to-peer network that requires transmission of location information to authenticate the audio device 10D. In this regard, adding authentication is consistent with Jannard’s goal of increasing the capabilities of a wearable audio device. *See, e.g.*, Jannard, [0009], [0010].

200. Thus, a person of ordinary skill in the art would have found it routine, straightforward and advantageous to apply Relan’s teachings of creating a peer-to-peer network based on GPS location to Jannard’s teachings of a wearable device (as modified by Wittenberg) with wireless peer-to-peer networking capabilities. A person of ordinary skill in the art would have known such a combination (yielding the claimed limitations) would predictably work and provide the expected functionality.

3. Claim 12

201. Claim [12] recites “The wearable device of claim 1, wherein the wearable device further comprises a GPS receiver and wherein the communication module forms a peer-to-peer network using location information from the GPS receiver.” See my discussion of claim [1], above (§IX.A). The limitation added by this claim is met by Jannard in view of Wittenberg and Relan.

202. Jannard discloses the wearable device further comprises a GPS receiver (*e.g.*, “a GPS receiver”). *See, e.g.*, Jannard:

- [0199]: “The sensor may measure a time-varying parameter, such as speed, acceleration, or jerk (*e.g.*, the rate of change in acceleration). The sensor may provide a signal indicative of an instantaneous or an average time-varying parameter at fixed, or at user-selectable intervals. For example, in one embodiment, the primary source 602 includes a **GPS receiver** and a clock and calculates velocity by dividing a measurement of distance traveled (as provided by the GPS receiver) by a measurement of time elapsed (as provided by the clock).”
- [0212] (“In one embodiment, the data port 608 is able to communicate with multiple source devices 602, 604....”)
- [0174] (“[T]he source device can be incorporated into the audio device 10C.”)
- [0267]: “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.**”

203. Jannard teaches a “GPS receiver” included in “primary source 602.” Jannard, [0199]. Jannard describes “primary source 602” as an example of “source

electronics,” which are integrated with the audio device. Jannard, [0174], [0212].

204. Relan discloses a communication module (e.g., “media exchange network 100,” including “home/location network 107”) forms a peer-to-peer network (e.g., “peer-to-peer networking capability”) using location information (e.g., “us[ing] location information”) from the GPS receiver (e.g., “GPS receiver 130”). See, e.g., Relan:

- [0049] (“In an embodiment according to the present invention, the media exchange network 100 may also comprise a plurality of media peripheral devices 118 including a mobile multi-media gateway 111 that may have **a global positioning system (GPS) receiver 130...**”)
- [0059] (“[T]he home/location network 107 may include at least one of home PNA (phoneline networking alliance) infrastructure, home cable infrastructure, Ethernet infrastructure, and wireless infrastructure, for example, providing **peer-to-peer networking capability within the first home/location 103. ...The terms peer-to-peer and peer2peer may be used herein, interchangeably, to refer to a communication network in which information is transferred or shared by one user with another, using the same program or type of program (i.e., peers).** Such networks typically operate without intermediate storage and retrieval of the information on a storage facility or server.”)[0084] (“In an embodiment of the present invention, a request for the exchange of multimedia information from a first MPS such as, for example, the MPS 102, and a second MPS such as, for example, the MPS 117 may employ security-related information from a user for authenticating the request. Such security-related information may, for example, comprise a password and/or a digital certificate provided by the user or by the MPS of the user requesting the exchange. In addition, an embodiment of the present invention may employ position or physical location information about the geographic location of a mobile terminal or handset such as, for example, a cellular phone that is associated with the user making the request. Other such devices may also be employed including, for example, two-way paging devices, mobile multimedia handsets, mobile multimedia gateways, personal digital assistants, and personal computers equipped

with wireless interfaces, to name only a few devices. Such password or digital certificate information, and geographic position/location information of an associated mobile handset may, for example, be evaluated and/or verified by the source or destination of a requested multimedia information exchange, before exchange of the multimedia information is permitted to proceed.”)

- [0085] (“In an embodiment of the present invention, the position/location information may comprise, for example, coordinate information from a global positioning system (GPS) receiver (or another positioning means) within or in communication with the mobile handset, and/or may comprise information identifying the location of the mobile handset within a wireless network infrastructure such as, for example, the serving cell and/or sector of a cellular telephone network. An example of one embodiment of such a mobile handset device is shown in FIG. 1 as mobile multimedia gateway 111 with GPS receiver 130.”)

205. Relan teaches a “mobile handset” comprising a “GPS receiver 130” with “media exchange network 100.” Relan, [0054], [0085], [0295]. The media exchange network 100 provides for “the exchange of multimedia information” and includes a “home/location network 107” that has “peer-to-peer networking capability.” Relan, [0049], [0059], [0085]. Relan further teaches using “information identifying the location” from the GPS receiver to authenticate the user when establishing the peer-to-peer connection. Relan, [0084]-[0085].

206. As described above in §IX.B.2, a person of ordinary skill in the art would have been motivated to apply Relan’s teaching of using a peer-to-peer network and using GPS to authenticate a user when establishing a peer-to-peer network to Jannard’s audio device.

C. GROUND 3: GROUND 1 IN FURTHER VIEW OF BARNES (Claim 15):

1. Overview of Barnes

207. Barnes discloses an e-commerce system that uses a “voice command that instructs the device” to make a “purchase” online. Barnes, [0124]. Barnes describes a “portable device 101” that “is preferably handheld, but may also be worn around the neck, on the hip, attached to the arm of the user, or carried in any convenient manner.” Barnes, [0034]. The user “input[s] a command (by saying ‘purchase’) to purchase the product,” leading to the device carrying out the “commercial transactions” in the “commerce module.” Barnes, [0124], [0128].

2. Motivation to Apply Barnes’ Teachings to Jannard and Wittenberg

208. Jannard, Wittenberg, and Barnes are analogous art. Jannard, Wittenberg, and Barnes are in the same field as the ’692 patent, which includes wearable voice-command responsive devices. Jannard, Wittenberg, and Barnes are reasonably pertinent to alleged problems addressed by the ’692 of increasing the capabilities of an audio device. E.g., ’692 patent, 1:22-34; Jannard, Abstract, [0189]; Barnes, [0033]-[0035] (“The present invention is a multi-function communications device 101 that includes conventional mobile phone and personal digital assistant (PDA) capabilities and preferably one or more of the functional modules shown in FIG. 1 such as....”); *see also* my discussion on the motivation to combine Wittenberg with Jannard, §IX.A.3. For example, Barnes’ device is “a

multi-function” audio-capable device that combines various “capabilities” of the “conventional mobile phone and personal digital assistant (PDA).” Barnes, [0033]. Such combining capacities of devices is reasonably pertinent to the problem addressed by the ’692 patent of combining the capacities of several devices into one audio device. ’692 patent, 1:35-40. This is also pertinent to the problem addressed by Jannard, relating to the control of additional functionality from a single device. Jannard, [0009]-[0010].

209. A person of ordinary skill in the art would have been motivated to apply Barnes’ teachings of “input [of] a command (by saying ‘purchase’) to purchase [a] product” to Jannard’s multi-faceted audio device (as modified by Wittenberg) because 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. Barnes, [0124]; Jannard, [0292], [0299], [0301]. Thus, a person of ordinary skill in the art would have been motivated to look to analogous art—such as Barnes—to identify applications of internet browsing using voice commands, and would have identified purchasing products as one such application.

210. Moreover, a person of ordinary skill in the art would have found it straightforward to use Jannard’s device to carry out online shopping. Jannard’s device already has voice commands for “internet browsing.” Jannard, [0299]. And

online shopping was a well-known form of internet browsing. *See, e.g.*, EX.1014 (Australian Patent Application No. 2005201979 A1), 1:21-22 (“With the advances in information communication technology, online shopping by means of Web pages on the Internet has now come into widespread use.”). For example, a person of ordinary skill in the art would have understood that, when applying Barnes’s teachings of online-shopping to Jannard’s audio device, the voice commands used to perform internet browsing would include commands to purchase a product. Feedback from the webpage would be provided through Jannard’s text-to-speech functionality that is used during internet browsing. Jannard, [0299], [0316].

211. Thus, a person of ordinary skill in the art would have found it routine, straightforward and advantageous to apply Barnes’ teachings of product purchases by voice command to Jannard’s teachings of a voice-command-responsive device (as modified by Wittenberg) 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 15

212. **Claim [15] recites “The wearable device of claim 1, wherein the initiating of operations comprises commercial transactional operations.”** See my discussion of claim [1], above (§IX.A). The limitation added by this claim is

met by Jannard in view of Barnes.

213. Jannard discloses the initiating of operations (*see* [1.f]) comprises commercial transactional operations (*e.g.*, “advertising” from the “internet”). *See, e.g.*, Jannard:

- [0299] (“The presently marketed BLACKBERRY™ devices can communicate with a variety of wireless networks for receiving email, phone calls, and/or **internet browsing**... 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**...”)
- [0301] (“**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... **advertising**, etc.”)

214. See my discussion of element [1.f], above (§IX.A).

215. Jannard teaches “internet browsing” through a “source device B” like a Blackberry. Jannard, [0299]. Jannard also teaches a “content source 302” “of any variety of information sources” providing “advertising,” which is a form of commercial activity. Jannard, [0301].

216. Barnes discloses initiating commercial transactional operations (*e.g.*, “input[ting] a command (by saying ‘purchase’)” in order to carry out “commercial transactions”). *See, e.g.*, Barnes:

- [0124] (“Alternately, the user can scan a product name or number and **input a command (by saying “purchase”) to purchase the product**. In response, the device 101 retrieves payment information, user information and an address (*e.g.*, web address) of a vender for such products from

memory, service registry or other source and transmits the product information, user information, and payment information to the vender ***to purchase the product.***")

- [0128] (***"The commerce module 135 facilitates commercial transactions and in particular, performs the steps of exchanging payment information, receiving acknowledgement of satisfactory completion of the transaction,*** providing an indication of the satisfactory or unsatisfactory completion of the transaction to the user, and processing and storing information ***resulting from the transaction*** (e.g., a hotel room number, a ticket number, etc.).")

217. Barnes discloses that the user can "input a command (by saying 'purchase') to purchase the product." Barnes, [0124]. And thus, the device carries out the "commercial transactions." Barnes, [0128].

218. And as I discussed in §IX.C.2, a person of ordinary skill in the art would have been motivated to apply Barnes' teachings of purchasing products using voice command to Jannard's device, as modified by Wittenberg.

X. SECONDARY CONSIDERATIONS

219. I have seen no evidence in the '692 patent's prosecution history or elsewhere supporting any secondary considerations arguments, or evidence of nexus to any of the Challenged Claims. See generally '692 File History. Indeed, per my discussion of the prior art referenced herein, any purported solutions to problems or unexpected results in the '692 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

220. 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 '692 patent obvious.

221. 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.

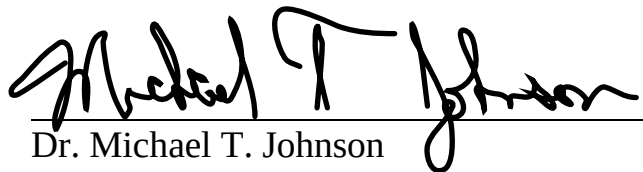
222. I reserve the right to supplement my opinions in the future to respond to any arguments that Patent Owner or its expert(s) may raise and to take into account new information as it becomes available to me.

223. 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.

224. 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.

225. 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.

226. Executed this 30 day of August 2024, in Lexington KY.



Dr. Michael T. Johnson

APPENDIX A

Curriculum Vitae

Michael T. Johnson

Associate Dean for Undergraduate Education and Student Success
Professor, Electrical and Computer Engineering
University of Kentucky College of Engineering
453 F. Paul Anderson Tower, Lexington, KY 40506, Phone (859) 257-0717
Email: mike.johnson@uky.edu
Web page: <http://johnson.engineering.uky.edu/>

CURRICULUM VITAE EXECUTIVE SUMMARY

Publications: **52 journal papers, over 150 refereed publications/presentations**
 Citations ≈3600, h-index 31, i10-index 70 (Google scholar Jan 2024)

Grant funding: **6 major grants, over \$3.5 million in external funding (\$2.7m as PI)**
 Multi-institution, interdisciplinary, international collaborations

Students: **8 PhD graduates, 14 MS graduates**

Leadership: **Associate Dean for Undergrad Education and Student Success, UK COE (2023-)**
 President, Southeast ECE Department Heads Association (2021-22)
 SEC Academic Leadership Development Program, (2019-2020)
 Department Chair, UK ECE (2016-2023)
 Director, UK CoE Scholars in Engineering Leadership (2017-2020)
 Chair of Marquette Board of Graduate Studies (2012-2016)
 Director of Graduate Studies, Marquette EECE (2009-2012)
 Senior Visiting Professor, Tsinghua University (2008-2009, 2014-2015)
 Engineering Manager (1993-1996)

Honors: **Featured in “Groundbreaking Thinkers in Wisconsin” series 2007**
 Marquette COE Researcher of the Year 2006-2007
 Engineers and Scientists of Milwaukee Young Engineer of the Year 2006
 HKN EECE Teacher of the year 2005, 2014

RESEARCH SPECIALIZATION

- Speech Processing: speech recognition and enhancement, natural language processing
- Signal Processing: bioacoustics, microphone array processing, signal enhancement
- Pattern Recognition and Machine Learning: statistical estimation and classification

EDUCATION

Doctor of Philosophy, School of Electrical and Computer Engineering, August 2000
Purdue University, West Lafayette, IN GPA: 4.00/4.00
"Incorporating Prosodic Information and Language Structure into Speech Recognition Systems"
Advisor: Dr. Leah H. Jamieson

Master of Science in Electrical Engineering (concentration Digital Signal Processing), Dec. 1994
University of Texas at San Antonio, San Antonio, TX GPA: 4.00/4.00
"An Adaptive Texture Model Based on Markov Random Fields"
Advisor: Dr. Mita Desai

Bachelor of Science, Engineering with Electrical Concentration, April 1990
LeTourneau University, Longview, TX GPA: 3.60/4.00

Bachelor of Science, Computer Science and Engineering with minor in Mathematics, April 1989
LeTourneau University, Longview, TX GPA: 3.57/4.00

ACADEMIC WORK EXPERIENCE

8/2016 – present University of Kentucky, Lexington, KY

- *Associate Dean for Undergraduate Education, College of Engineering* (2023-present)
- *Department Chair, Electrical and Computer Engineering* (2016-2023)
- *Full Professor, Electrical and Computer Engineering* (2016 - present)

8/2000 – 8/2016 Marquette University, Milwaukee, WI

- *Full Professor, Electrical and Computer Engineering*: (2013 - 2016)
- *Director of Graduate Studies, Electrical and Computer Engineering*: (2009 - 2012)
- *Associate Professor with Tenure, Electrical and Computer Engineering* (2007 – p2009)
- *Assistant Professor, Electrical and Computer Engineering*: (2000 – 2007)

2008-2009 and 2014-2015 Tsinghua University, Beijing, China (two 1-year sabbaticals)

- *Senior Visiting Faculty, Electronic Engineering Dept.* While on sabbatical, taught courses and collaborated on research in speech processing with Prof. Liu Jia.

9/1996 – 8/2000 Purdue University, West Lafayette, IN

- *Graduate Research Assistant, Audiology Dept.*: (5/97 to 5/00) Research focused on studying the consistency of speech articulator movement in children and young adults.
- *Graduate Research Assistant, ECE Dept.*: (1/97 to 1/98) Part of an NSF-funded research grant working on interfacing speech recognition and language processing systems.
- *Teaching Assistant, ECE Dept.*: (9/96 to 12/96 and 1/98 to 5/98) Assisted with a course in Speech Processing and taught a laboratory course in Digital Circuits.

1/1993 - 9/1993 University of Texas at San Antonio, San Antonio, TX

- *Graduate Research Assistant, UTSA Image Processing Lab*: Did research in the field of biological image processing as part of masters thesis in mammogram image analysis.

INDUSTRIAL WORK EXPERIENCE

9/1993 - 7/1996 SNC Manufacturing, Oshkosh, WI

- *Engineering Manager, Telecommunications Division (11/94 - 7/96)*: Responsible for all new and existing product lines, including design, project management, and personnel.
- *Senior Engineer (9/93-11/94)*: Project lead for new products, primarily in the area of SNC's Lyte Lynx line of fiber-optic interfaces for high-voltage environments.

9/1991 - 1/1993 Datapoint Corporation , San Antonio, TX

- *Engineer II*: Hardware engineer on the ArcnetPlus networking and MINX videoconferencing projects. Lead engineer on the PC HUB project for ArcnetPlus.

5/1990 - 9/1991 Micronyx, Inc. / Micro Technology Services, Inc., Richardson, TX

- *Design Engineer*: Telecommunications and embedded microprocessors design, at both hardware and software levels, for Micronyx and its spin-off company MTSI.

5/1989 - 8/1989 GTE , Ft. Wayne, IN *Engineering Intern*: Designed circuits for Business Services, including a project restructuring an analog phone system into a digital carrier system.

6/1988 - 8/1988 NBS , Ft. Wayne, IN *Electrical Technician*: Gained experience with troubleshooting circuit boards, soldering, and designing and building test equipment.

PUBLICATIONS

Peer-reviewed Journal Publications

1. Rahim Soleymanpour, Mohammad Soleymanpour, Anthony J Brammer, Michael T Johnson, Insoo Kim, Speech Enhancement Algorithm Based on a Convolutional Neural Network Reconstruction of the Temporal Envelope of Speech in Noisy Environments, IEEE Access, Volume 11, 5328-5336, January 2023.
2. Julie N. Oswald, Amy M. Van Cise, Angela Dassow, Taffeta Elliott, Michael T. Johnson, Andrea Ravignani, Jeffrey Podos, A Collection of Best Practices for the Collection and Analysis of Bioacoustic Data, Applied Sciences, Volume 12, No. 23, November 2022.
3. Chao Li; Wang, Qiyue Wang, Wenhua Jiao, Michael T. Johnson, Yuming Zhang, Deep Learning-Based Detection of Penetration from Weld Pool Reflection Images, Volume 99 No. 2, Welding Journal, September 2020.
4. Qiyue Wang, Wenhua Jiao, Rui Yu; Michael T. Johnson, Yuming Zhang, Virtual Reality Robot-Assisted Welding Based on Human Intention Recognition, IEEE Transactions on Automation Science and Engineering, Volume 17 No. 2, April 2020.
5. Yi Liu; Liang He; Jia Liu; Michael T. Johnson, Introducing phonetic information to speaker embedding for speaker verification, EURASIP Journal on Audio, Speech, and Music Processing, December 2019.
6. Qiyue Wang, Yongchao Cheng, Wenhua Jiao, Michael T. Johnson, Yuming Zhang, Virtual reality human-robot collaborative welding: A case study of weaving gas tungsten arc welding, Journal of Manufacturing Processes, Volume 49, 2019.
7. Qiyue Wang, Wenhua Jiao, Rui Yu, Michael Johnson, YuMing Zhang, 2019. Modeling of Human Welder Operation in Virtual Reality Human-Robot Interaction Welding System. IEEE Robotics and Automation Letters, 4(3): 2958-2964, July 2019.
8. Ruijie Yan, Liangrui Peng, Shanyu Xiao, Michael T. Johnson, Shengjin Wang, Dynamic temporal residual network for sequence modeling. International Journal on Document Analysis and Recognition, Volume 22, 235–246, July 2019.
9. Liang He, Xianhong Chen, Can Xu, Yi Liu, Jia Liu, Michael T. Johnson, Latent class model with application to speaker diarization, EURASIP Journal on Audio, Speech, and Music Processing, July 2019.
10. Jian Kang, Wei-Qiang Zhang, Weiwei Liu, Jia Liu., Michael T. Johnson, Lattice Based Transcription Loss for End-to-End Speech Recognition, Journal of Signal Processing Systems, Volume 90 Issue 7, pp 1013-1023, July 2018
11. Jian Kang, Wei-Qiang Zhang, WeiWei Liu, Jia Liu, Michael T. Johnson, Advanced recurrent network-based hybrid acoustic models for low resource speech recognition, EURASIP Journal on Audio, Speech, and Music Processing, 2018.
12. Liang He, Xianhong Chen, Can Xu, Jia Liu, Michael T. Johnson, Local Pairwise Linear Discriminant Analysis for Speaker Verification, IEEE Signal Processing Letters Volume 25 No. 10, 1575-1579, 2018.
13. Jeffrey Berry, Andrew Kolb, James Schroeder, Michael T. Johnson, Jaw Rotation in Dysarthria Measured With a Single Electromagnetic Articulography Sensor, American Journal of Speech – Language Pathology, June 2017.
14. Ji, An; Johnson, Michael T; Berry, Jeffrey J., Parallel Reference Speaker Weighting for Kinematic-Independent Acoustic-to-Articulatory Inversion, IEEE/ACM Transactions on Audio, Speech, and Language Processing",24,10, October 2016, 1865-1875.

15. Xu-Kui Yang, Liang He, Dan Qu, Wei-Qiang Zhang, Michael T Johnson, Semi-supervised feature selection for audio classification based on constraint compensated Laplacian score, *EURASIP Journal on Audio, Speech, and Music Processing*, 1, 2016, 1-10.
16. Liu Wei-Wei, Cai Meng, Zhang Wei-Qiang Zhang, Liu Jia, Johnson Michael T., "Discriminative Boosting Algorithm for Diversified Front-End Phonotactic Language Recognition," *Journal of Signal Processing Systems*, February 2016.
17. Wei-Qiang Zhang, Cong Guo, Qiao Zhang, Jian Kang, Liang He, Jia Liu, and Michael T. Johnson, A Speech Enhancement Algorithm Based on Computational Auditory Scene Analysis. *Journal of Tianjin University*, in press, 2015. (Chinese journal. Original language citation: 张卫强, 郭聪, 张乔, 康健, 何亮, 刘加, and Michael T. Johnson, 种基于计算听觉场景分析的语音增强算法, 天津大学学报.)
18. B.T.W. Bochera, K. Cherukurib, J.S. Makib, M. Johnson, D.H. Zitomer, Relating methanogen community structure and anaerobic digester function, *Water Research*, 70 (1), March 2015, 425-435.
19. PM Scheifele, MT Johnson, M Fry, B Hamel, K Laclede, Vocal classification of vocalizations of a pair of Asian Small-Clawed otters to determine stress, *The Journal of the Acoustical Society of America*, 138 (1), EL105-EL109, 2015.
20. Trawicki, Marek B. and Johnson, Michael T., "Beta-order minimum mean-square error multichannel spectral amplitude estimation for speech enhancement", *International Journal of Adaptive Control and Signal Processing*, January 2015.
21. Arik Kershenbaum, Daniel Blumstein, Marie Roch, Michael T. Johnson, et. al., Acoustic sequences in non-human animals: a tutorial review and prospectus, *Biological Reviews*, 2014.
22. C Yu, KK Wójcicki, PC Loizou, JHL Hansen, MT Johnson, "Evaluation of the importance of time-frequency contributions to speech intelligibility in noise", *The Journal of the Acoustical Society of America*, vol. 135, no. 5, May 2014, 3007-3016.
23. Marek B. Trawicki, Michael T. Johnson, "Speech enhancement using Bayesian estimators of the perceptually-motivated short-time spectral amplitude (STSA) with Chi speech priors", *Speech Communication*, vol. 57, no. 2, February 2014, pp101-103.
24. Liu Weiwei, Zhang Weiqiang, Johnson Michael T., Liu Jia, "Homogenous ensemble phonotactic language recognition based on SVM supervector reconstruction", *EURASIP Journal on Audio, Speech, and Music Processing* vol. 2014 no. 1, January 2014, pp 1-13.
25. Junhong Zhao, Wei-Qiang Zhang, Hua Yuan, Michael T Johnson, Jia Liu, Shanhong Xia, "Exploiting contextual information for prosodic event detection using auto-context", *EURASIP Journal on Audio, Speech, and Music Processing*, vol. 2013, no. 1, December 2013 pp 1-14.
26. Marek B. Trawicki, Michael T. Johnson, "Distributed multichannel speech enhancement based on perceptually-motivated Bayesian estimators of the spectral amplitude", *IET Signal Processing*, vol. 7, no.4, April 2013, pp. 337-344.
27. An Ji, Michael T. Johnson, Edward J. Walsh, JoAnn McGee, Doug L. Armstrong, Discrimination of individual tigers (*Panthera tigris*) from long distance roars, *The Journal of the Acoustical Society of America*, vol. 133 no. 3, March 2013, pp1762-1769.
28. Yongzhe Shi, Weiqiang Zhang, Jia Liu, Michael T. Johnson, "RNN language model with word clustering and class-based output layer", *EURASIP Journal on Audio, Speech, and Music Processing* vol. 2013 no. 1, January 2013, pp1-7.

29. Peter M. Scheifele, Michael T. Johnson, David C. Byrne, John G. Clark, Ashley Vandlik, Laura W. Kretschmer, Kristine E. Sonstrom, "Noise impacts from professional dog grooming forced-air dryers", *Noise and Health*, vol. 14 no. 60, October 2012, p224-226.
30. Wen-Lin Zhang, Wei-Qiang Zhang, Bi-Cheng Li, Dan Qu and Michael T. Johnson, "Bayesian Speaker Adaptation Based on a New Hierarchical Probabilistic Model", *IEEE Transactions on Speech and Language Processing*, vol. 20 no. 7, July 2012, pp2002-2015.
31. Yuxiang Shan, Yan Deng, Jia Liu, Michael T. Johnson, "Phone lattice reconstruction for embedded language recognition in LVCSR", *EURASIP Journal on Audio Speech and Music Processing*, vol. 2012, no. 15, April 2012, pp1-13.
32. Marek B. Trawicki, Michael T. Johnson, "Distributed multichannel speech enhancement with minimum mean-square error short-time spectral amplitude, log-spectral amplitude, and spectral phase estimation". *Signal Processing*, vol. 92 no. 2, February 2012, pp 345-356.
33. Peter M. Scheifele, Michael T. Johnson, Laura W. Kretschmer, John G. Clark, Deborah Kemper, Gopu Potty, "Ambient habitat noise and vibration at the Georgia Aquarium", *The Journal of the Acoustical Society of America*, vol. 132 no. 2, February 2012, EL88-EL94.
34. Wei Qiang Zhang, Liang He, Yan Deng, Jia Liu, Michael T. Johnson, "Time-Frequency Cepstral Features and Heteroscedastic Linear Discriminant Analysis for Language Recognition", *IEEE transactions on audio, speech, and language processing* vol. 19 no. 2, February 2011, pp.266-276.
35. Kuntoro Adi, Michael T. Johnson, and Tomasz S. Osiejuk, "Acoustic Censusing using Automatic Vocalization Classification and Identity Recognition," *Journal of the Acoustical Society of America*, vol. 127, no. 2, February 2010, pp 874-883.
36. Yao Ren, Michael T. Johnson, Patrick J. Clemins, Michael Darre, Sharon Stuart Glaeser, Tomasz S. Osiejuk, and Ebenezer Out-Nyarko, "A Framework for Bioacoustic Vocalization Analysis Using Hidden Markov Models", *Algorithms*, vol. 2 no. 3, November 2009, pp 1410-1428.s
37. Yanmin Qian, Jia Liu, Michael T. Johnson "Efficient Embedded Speech Recognition for Very Large Vocabulary Mandarin Car-Navigation Systems," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 3, August 2009, 1496-1500.
38. Yao Ren, Michael T. Johnson, Jidong Tao, "Perceptually motivated wavelet packet transform for bioacoustic signal enhancement", *Journal of the Acoustical Society of America*, vol. 124, no. 1, July 2008, 316-327.
39. Jidong Tao, Michael T. Johnson, Tomasz S. Osiejuk, "Acoustic model adaptation for ortolan bunting (*Emberiza hortulana* L.) song type classification", *Journal of the Acoustical Society of America*, vol. 123, no. 3, March 2008, 1582-1590.
40. Michael T. Johnson, Xiaolong Yuan, Yao Ren, "Speech signal enhancement through adaptive wavelet thresholding", *Speech Communication*, vol. 49, No. 2, February 2007, pp 123-133.
41. Patrick Clemins and Michael T. Johnson, "Generalized perceptual linear prediction (gPLP) features for animal vocalization analysis", *Journal of the Acoustical Society of America*, Vol. 120, No. 1, July 2006, pp 527-534.

42. Richard J. Povinelli, Michael T. Johnson, Andrew C. Lindgren, Felice Roberts, and Jinjin Ye, "Statistical Models of Reconstructed Phase Spaces for Signal Classification", *IEEE Transactions on Signal Processing*, vol. 54, no. 6, June 2006, 2178-2186.
43. Michael T. Johnson and Richard J. Povinelli, "Generalized phase space projection for nonlinear noise reduction", *Physica D*, No. 201, February 2005, pp 306-317.
44. Kevin M. Indrebo, Richard J. Povinelli, and Michael T. Johnson, "Sub-banded Reconstructed Phase Spaces for Speech Recognition", *Speech Communication*, Vol. 48, No. 7, July 2006, pp 760-774.
45. Michael T. Johnson, "Capacity and Complexity of HMM Duration Modeling Techniques", *IEEE Signal Processing letters*, Volume 12, No. 5, May 2005, pp 407-410.
46. Patrick J. Clemins, Michael T. Johnson, Kirsten M. Leong, and Anne Savage, "Automatic Classification and Speaker Identification of African Elephant (*Loxodonta Africana*) vocalizations", *Journal of the Acoustical Society of America*, Volume 117, No. 2, February 2005, pp 956-963.
47. Michael T. Johnson, Richard J. Povinelli, Jinjin Ye, Xiaolin Liu, Andrew Lindgren, and Kevin Indrebo, "Time-Domain Isolated Phoneme Classification using Reconstructed Phase Spaces", *IEEE Transactions on Speech and Audio Processing*, Vol. 13, No. 4, July 2005, pp 458-466.
48. Richard J. Povinelli, Michael T. Johnson, Andrew C. Lindgren, and Jinjin Ye, "Time Series Classification using Gaussian Mixture Models of Reconstructed Phase Spaces", *IEEE Transactions on Knowledge and Data Engineering*, vol. 16, no. 6, June 2004, 779-783.
49. Kelly Stonger and Michael T. Johnson, "Optimal Calibration of PET Crystal Position Maps using Gaussian Mixture Models", *IEEE Transactions on Nuclear Science*, Vol. 51, No. 1, 2004, pp 85-90.
50. Yang Liu, Mary P. Harper, Michael T. Johnson, and Leah H. Jamieson, "The Effect of Pruning and Compression on Graphical Representations of the Output of a Speech Recognizer", *Computer Speech and Language*, Volume 17, 2003, pp. 329-256.
51. A.M. Surprenant, S.L. Hura, M.P. Harper, L.H. Jamieson, G.Long, S.M. Thede, A. Rout, T.H. Hsueh, S.A. Hockema, M.T. Johnson, P. Srinivasan, and C.M. White, "Familiarity and Pronouncibility of Nouns and Names", *Behavior Research, Methods, Instruments, & Computers*, Vol. 31, Nov. 1999, pp. 638-649.
52. Anne Smith, Michael Johnson, Clare McGillem, and Lisa Goffman "On the Assessment of Stability and Patterning of Speech Movements", *Journal of Speech, Language, and Hearing Research*, Vol. 43, 2000, pp 277-286.

Peer-reviewed Conference Publications and Presentations

53. Mohammad Soleymanpour, Michael T. Johnson, Rahim Soleymanpour, Jeffrey Berry, Synthesizing Dysarthric Speech Using Multi-speaker TTS for Dysarthric Speech Recognition, International Conference on Acoustics, Speech, and Signal Processing (ICASSP), May 2022, Singapore.
54. Narjes Bozorg, Michael T Johnson, Mohammad Soleymanpour, Autoregressive Articulatory WaveNet Flow for Speaker-Independent Acoustic-to-Articulatory Inversion,

- 2021 international Conference on Speech Technology and Human-Computer Dialog (SpeD), pp. 156-161, October 2021.
55. Subash Khanal, Michael T Johnson, Mohammad Soleymanpour, Narjes Bozorg, Mispronunciation Detection and Diagnosis for Mandarin Accented English Speech, 2021 international Conference on Speech Technology and Human-Computer Dialog (SpeD), pp. 62-67, October 2021.
 56. Mohammad Soleymanpour, Michael T Johnson, Jeffrey Berry, Dysarthric Speech Augmentation Using Prosodic Transformation and Masking for Subword End-to-end ASR, 2021 international Conference on Speech Technology and Human-Computer Dialog (SpeD), pp. 42-46, October 2021.
 57. Mohammad Soleymanpour, Michael T Johnson, Jeffrey Berry, Comparison in Suprasegmental Characteristics between Typical and Dysarthric Talkers at Varying Severity Levels, 2021 international Conference on Speech Technology and Human-Computer Dialog (SpeD), pp. 52-56, October 2021.
 58. Mohammad Soleymanpour, Michael T Johnson, Jeffrey Berry, Increasing the Precision of Dysarthric Speech Intelligibility and Severity Level Estimate, International Conference on Speech and Computer, pp. 670-679, September 2021.
 59. Subash Khanal, Michael T. Johnson, Narjes Bozorg, Articulatory Comparison of L1 and L2 Speech for Mispronunciation Diagnosis, IEEE Spoken Language Technology Workshop (SLT) 2021, January 2021, Shenzhen China.
 60. Narjes Bozorg and Michael Johnson, Acoustic-to-Articulatory Inversion with Deep Autoregressive Articulatory-WaveNet, Interspeech 2020, October 2020, Shanghai China.
 61. Jiaqi Xie, Ruijie Yan, Shanyu Xiao, Liangrui Peng, Michael T. Johnson, Weiqiang Zhang, Dynamic Temporal Residual Learning for Speech Recognition, International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2020, May 2020, Barcelona, Spain.
 62. Jeffrey Berry and Michael T. Johnson, Perceived Acoustic Working Space Modulates Sensorimotor Learning, Motor Speech Conference, Santa Barbara, California, February 2020.
 63. Narjes Bozorg and Michael T. Johnson, Reference speaker selection for kinematic-independent acoustic-to-articulatory-inversion, J. Acoust. Soc. Am., vol. 145, no. 3, pp. 1932–1932, Mar. 2019, presentation at the Spring 2019 Meetings of the Acoustical Society of America.
 64. McGowan, Kevin B; Johnson, Michael T; Combs, Aleah; Soleymanpour, Mohammad; Acoustic, non-invasive measurement of velopharyngeal aperture using a high frequency tone, presentation at the Spring 2019 Meetings of the Acoustical Society of America.
 65. Narjes Bozorg, Michael T. Johnson, Jeffrey Berry, Comparing Articulography Consistency Between L1 and L2 Speakers, IEEE International Symposium on Signal Processing and Information Technology, Ajman, United Arab Emirates, December 2019.
 66. Narjes Bozorg, Michael T. Johnson, MLLR-PRSW for Kinematic-Independent Acoustic-to-Articulatory Inversion, IEEE International Symposium on Signal Processing and Information Technology, Ajman, United Arab Emirates, December 2019.
 67. Kevin McGowan, Michael Johnson, Aleah Combs, Mohammad Soleymanpour, Acoustic, non-invasive measurement of velopharyngeal aperture using a high frequency tone, Annual Spring Meeting of the Acoustical Society of America, Louisville, Kentucky, May 2019. (JASA Vol.145(3), pp.1931-1931.)

68. Kevin McGowan, Michael Johnson, Aleah Combs, Mohammad Soleymanpour, Acoustic, non-invasive measurement of velopharyngeal aperture using a high frequency tone, International Congress of Phonetic Sciences (ICPhS) 2019, Melbourne, August 2019.
69. Narjes Bozorg and Michael T. Johnson, Comparing Performance of Acoustic-to-Articulatory Inversion for Mandarin Accented English and American English Speakers, International Symposium on Signal Processing and Information Technology, Louisville, KY, December 2018.
70. Sidrah Liaqat, Narjes Bozorg, Neenu Jose, Patrick Conrey, Antony Tamasi and Michael T. Johnson, Domain Tuning Methods for Bird Audio Detection, Detection and Classification of Acoustic Scenes and Events 2018, Surrey, UK, November 2018.
71. Chao Li, Qiyue Wang, Jinsong Chen, Michael Johnson, Qiang Chang, YuMing Zhang, Machine Learning Based Detection of Weld Joint Penetration from Weld Pool Reflection Images, 14th IEEE International Conference on Automation Science and Engineering, Munich, Germany, 2018.
72. Yi Liu, Liang He, Jia Liu, Michael T. Johnson, Speaker Embedding Extraction with Phonetic Information, Interspeech Hyderabad, India, 2018.
73. Yi Liu, Liang He, Yao Tian, Zhuzio Chen, Jia Liu, Michael T. Johnson, Comparison of multiple features and modeling methods for text-dependent speaker verification, Automatic Speech Recognition and Understanding Workshop (ASRU), Okinawa, Japan, 2017.
74. Yi Liu, Liang He, Weiqiang Zhang, Jia Liu, Michael T. Johnson, Investigation of Frame Alignments for GMM-based Digit-prompted Speaker Verification, Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, Honolulu, Hawaii, November 2018.
75. Weilian Song, Tawfiq Salem, Nathan Jacobs, Michael T. Johnson, Detecting the Presence of Bird Vocalizations in Audio Segments Using a Convolutional Neural Network Architecture, 4th International Symposium on Acoustic Communication by Animals, Omaha, NE, July 2017.
76. Peter M. Scheifele, Giulia Raponi, Haylea Roark, and Lesa Scheifele, Michael T. Johnson, Neenu Jose, Vocalization and Hearing Acuity of the Asian Small-Clawed Otter (*Amblonyx cinereus*) in Response to Anthropogenic Noise, 4th International Symposium on Acoustic Communication by Animals, Omaha, NE, July 2017.
77. Jeff Berry, Andrew Kolb, James Schroeder, Michael T Johnson, Jaw rotation in dysarthria measured with a single electromagnetic articulography sensor, Proceedings of Spring 2017 Meetings on Acoustics, Boston, June 2017
78. Jeffrey J Berry, Abigail Stoll, Deriq Jones, Seyedramin Alikiaamiri, Michael T Johnson, Second language pronunciation training using acoustic-to-articulatory inversion, Proceedings of Spring 2017 Meetings on Acoustics, Boston, June 2017.
79. Berry, Jeffrey J; Bagin, Ramie; Schroeder, James; Johnson, Michael T; Sensitivity and specificity of auditory feedback driven articulatory learning in virtual speech, Proceedings of Spring 2017 Meetings on Acoustics, Boston, June 2017.
80. Andrew J Kolb, Michael T Johnson, Jeffrey Berry, Interpolation of Tongue Fleshpoint Kinematics from Combined EMA Position and Orientation Data, Sixteenth Annual Conference of the International Speech Communication Association (Interspeech), September 2015.

81. Yi Liu, Yao Tian, Liang He, Jia Liu, Michael T Johnson, Simultaneous Utilization of Spectral Magnitude and Phase Information to Extract Supervectors for Speaker Verification, ASV Spoofing Challenge 2015, presented at Interspeech 2015, September 2015.
82. Gui Jin, Michael T Johnson, Jia Liu, Xiaokang Lin, Voice conversion based on Gaussian mixture modules with Minimum Distance Spectral Mapping, 2015 5th International Conference on Information Science and Technology (ICIST), April 2015, pp. 356-359.
83. Berry, Jeffrey, Cassandra North, and Michael T. Johnson, Dynamic aspects of articulating with a virtual vocal tract in dysarthria, International Conference on Motor Speech, 2014.
84. Berry, Jeffrey, A Kolb, C North, and Michael T. Johnson, Acoustic and Kinematic Characteristics of Vowel Production through a Virtual Vocal Tract in Dysarthria, Interspeech, Singapore, September 2014.
85. Berry, Jeffrey, John Jaeger, M Wiedenhoeft, B Bernal, and Michael T. Johnson, Consonant Context Effects on Vowel Sensorimotor Adaptation, Interspeech, Singapore, September 2014.
86. Ji, An, Michael T. Johnson, and Jeffrey Berry, Palate-referenced Articulatory Features for Acoustic-to-Articulator Inversion, Interspeech, Singapore, September 2014.
87. Jianglin Wang and Michael T. Johnson, "Physiologically-motivated Feature Extraction for Speaker Identification", International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2014, Florence, Italy.
88. Jeffrey Berry, Cassandra North, and Michael T. Johnson, "Sensorimotor adaptation of speech using real-time articulatory resynthesis", International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2014, Florence, Italy.
89. An Ji, Michael T. Johnson, and Jeffrey Berry, "The Electromagnetic Articulography Mandarin Accented English (EMA-MAE) Corpus of Acoustic and 3D Articulatory Kinematic Data", International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2014, Florence, Italy.
90. Jeffrey Berry, An Ji, and Michael T. Johnson, "Dynamic aspects of vowel production in Mandarin-accented English", American Speech-Language-Hearing Association 2013, Chicago, IL, November 2013.
91. Jeffrey Berry, Christine Belchel, Cassandra North, and Michael T. Johnson, "Learning novel articulatory-acoustic mappings in dysarthria", American Speech Language Hearing Association 2013, Chicago, IL, November 2013.
92. Jianglin Wang and Michael T. Johnson, "Vocal source features for bilingual speaker identification", China SIP, July 2013, Beijing China.
93. An Ji, Michael T. Johnson, Jeffrey Berry, "Articulatory space calibration in 3D Electromagnetic Articulography", China SIP, July 2013, Beijing China.
94. Jeffrey Berry, Cassandra North, Benjamin Meyers, Michael T. Johnson, Speech sensorimotor learning through a virtual vocal tract, Proceedings of Spring 2013 Meetings on Acoustics, Montreal, June 2013.
95. An Ji, Jeffrey Berry, Michael T. Johnson, Vowel production in Mandarin accented English and American English: Kinematic and acoustic data from the Marquette

University Mandarin accented English corpus, Proceedings of Spring 2013 Meetings on Acoustics, Montreal, June 2013.

96. Jianglin Wang, MT Johnson, "Residual Phase Cepstrum Coefficients with Application to Cross-lingual Speaker Verification", Interspeech 2012, Portland, September 2012.
97. Trawicki, Marek B. and Michael T. Johnson. "Improvements of the Beta-Order Minimum Mean-Square Error (MMSE) Spectral Amplitude Estimator using Chi Priors," Interspeech 2012, Portland, September 2012.
98. An Ji, Michael T. Johnson, Jeffrey Berry, "Tracking articulator movements using orientation measurements", 2012 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, China, July 2012.
99. Marek B. Trawicki, Michael T. Johnson, An Ji, Tomasz S. Osiejuk, "Multichannel speech recognition using distributed microphone signal fusion strategies", 2012 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, China, July 2012.
100. Jianglin Wang, An Ji, Michael T. Johnson, "Features for phoneme independent speaker identification", 2012 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, China, July 2012.
101. Michael T. Johnson and Patrick Clemins, "Individual Identification using Hidden Markov Models for Population Monitoring and Assessment", Invited Presentation to Special Session "Topical Meeting on Signal Processing of Subtle and Complex Acoustic Signals in Animal Communication", Acoustical Society of America Spring 2010 Meeting, Baltimore, Maryland, April 2010.
102. Marek Trawicki and Michael T. Johnson, "Optimal Distributed Microphone Phase Estimation," International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2009, Taiwan, April 2009, pp 2177-2180.
103. Yao Ren and Michael T. Johnson, "Auditory Coding based Speech Enhancement", International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2009, Taiwan, April 2009, pp 4685-4688.
104. Marek B. Trawicki, Yao Ren, Michael T. Johnson, Tomasz S. Osiejuk, Distributed Multi-Channel Speech Enhancement of Ortolan Bunting (*Emberiza Hortulana*) Vocalizations, Acoustic Communication by Animals, Corvallis, Oregon, 2008, pp. 276-277.
105. Kuntoro Adi, Michael T. Johnson, Tomasz S. Osiejuk Hidden Markov Model (HMM) based animal acoustic censusing, Acoustic Communication by Animals, Corvallis, Oregon, 2008, pp. 3-4.
106. Ebenezer Otu-Nyarko, Peter Scheifele, Michael Darre, Michael Johnson, Classification of stressful vocalizations of captive laying chickens using the Hidden Markov Model (HMM), Acoustic Communication by Animals, Corvallis, Oregon, 2008, pp. 176-177.
107. Kristine Sonstrom, Peter Scheifele, Michael Darre, Michael Johnson, The classification of vocalizations to identify social groups of beluga whales in the St. Lawrence River Estuary using the Hidden Markov Model, Acoustic Communication by Animals, Corvallis, Oregon, 2008, pp 251-252.

108. Kuntoro Adi, Kristine E. Sonstrom, Peter M. Scheifele, Michael T. Johnson, "Unsupervised validity measures for vocalization clustering," International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2008, April 2008, pp 4377-4380.
109. Yao Ren, Michael T. Johnson, "An Improved SNR estimator for speech enhancement," International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2008, April 2008, pp 4901-4904.
110. Mohamed A. Mneimneh, Michael T. Johnson, Richard J. Povinelli, "A Heart Cell Model for the Identification of Myocardial Ischemia," International Conference on Health Informatics, Funchal Madeira, Portugal, January 2008.
111. Z. W. Slavens, R. S. Hinks, J. A. Polzin, and M. T. Johnson, "Improved MR Image Magnification by Generalized Interpolation of Complex Data," Proceedings of the 15th Annual Meeting of ISMRM, Berlin, Germany, (Abstract #1887), May 2007.
112. Craig A. Struble, Richard J. Povinelli, Michael T. Johnson, Dina Berchanskiy, Jidong Tao, Marek Trawicki, "Combined Conditional Random Fields and n-Gram Language Models for Gene Mention Recognition, Proceedings of the Second BioCreative Challenge Evaluation Workshop, Madrid, Spain, April 2007.
113. Xi Li, Jidong Tao, Michael T. Johnson, Joseph Soltis, Anne Savage, Kirsten M. Leong,, John D. Newman, Stress and Emotion Classification using Jitter and Shimmer Features, International Conference on Acoustics Speech and Signal Processing 2007 (ICASSP07), Honolulu, Hawaii, April 2007, pp 1081-1084.
114. Mohamed A. Mneimneh, Richard J. Povinelli, Michael T. Johnson " An Integrative Approach for the Measurement of QT interval" Computers in Cardiology, Valencia, Spain, September 2006, pp 329-332.
115. Mohamed A. Mneimneh, Edwin E. Yaz, Michael T. Johnson, Richard J. Povinelli "Adaptive Kalman Filter Approach for the Baseline Removing Baseline Wandering " Computers in Cardiology, Valencia, Spain, September 2006, pp. 253-256.
116. Richard J. Povinelli, Mohamed A. Mneimneh, Michael T. Johnson "Cardiac Model Based Approach to QT Estimation" Computers in Cardiology, Valencia, Spain, September 2006, pp. 333-336.
117. Patrick Clemens, Marek B. Trawicki, Kuntoro Adi, Jidong Tau, and Michael T. Johnson, "Generalized Perceptual Features for Vocalization Analysis across Multiple Species", International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2006, Toulouse, France, May 2006.
118. Kuntoro Adi, Michael T. Johnson, Cepstral moment normalization for robust individual identification of ortolan bunting (*emberiza hortulana L*), presented at the Spring 2006 meetings of the Acoustical Society of America, May 2006, Providence, RI
119. Anthony D. Ricke, Richard J. Povinelli, Michael T. Johnson. "Segmenting Heart Sound Signals," Computers in Cardiology, Leon, France, September 2005.
120. Marek Trawicki and Michael T. Johnson, "Automatic Song-Type Classification and Speaker Identification of Norwegian Ortolan Bunting (*Emberiza Hortulana*) Vocalizations", IEEE International Conference on Machine Learning in Signal Processing (MLSP), Mystic, Connecticut, September 2005.

121. Patrick J. Clemens and Michael T. Johnson, "Unsupervised Classification of Beluga Whale Vocalizations", Spring 2005 Meetings of the Acoustical Society of America, Vancouver, May 2005.
122. Kevin M. Indrebo, Richard J. Povinelli, Michael T. Johnson. "Third-Order Moments of Filtered Speech Signals For Robust Speech Recognition," International Conference on Non-Linear Speech Processing (NOLISP) 2005, Barcelona, Spain, 151-157.
123. Jidong Tao and Michael T. Johnson, "Comparison and Evaluation of Animal Vocalization Enhancement Techniques", Fall 2004 Meetings of the Acoustical Society of America, San Diego, November 2004.
124. Kuntoro Adi and Michael T. Johnson, "Automatic Song-type classification and individual identification of Ortolan Bunting (*Emberiza Hortulana* L) vocalizations", Fall 2004 Meetings of the Acoustical Society of America, San Diego, November 2004.
125. Heather E. Ewalt and Michael T. Johnson, "Combining Multi-source Wiener Filtering with Parallel Beamformers to Reduce Noise from Interfering Talkers", International Conference on Signal Processing (ICSP) 2004, Beijing, August 2004.
126. Kevin M. Indrebo, Richard J. Povinelli, and Michael T. Johnson, "A comparison of reconstructed phase spaces and cepstral coefficients for multi-band phoneme classification", ICSP 2004, Beijing, August 2004.
127. Patrick Clemens and Michael T. Johnson, "Generalized perceptual features for animal vocalization classification", Spring 2004 Meetings of the Acoustical Society of America, New York, May 2004.
128. Andrew Lindgren, Michael T. Johnson, and Richard J. Povinelli, Joint Frequency Domain and Reconstructed Phase Space Features for Speech Recognition", International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2004, Montreal, May 2004.
129. Mike Zimmerman, Richard J. Povinelli, Michael T. Johnson, and Kris Ropella, "A Reconstructed Phase Space Approach for Distinguishing Ischemic from Non-Ischemic ST Changes using Holter ECG Data", Computers in Cardiology, Thessalonica, Greece, September, 2003.
130. Patrick Clemens and Michael T. Johnson, "Automatic Classification of African Elephant (*Loxodonta africana*) Follicular and Luteal Rumbles", 1st International Conference on Acoustic Communication by Animals, Baltimore, Maryland, July 2003, pp 81-82.
131. Patrick Clemens and Michael T. Johnson, "Automatic Type Classification and Speaker Identification of African Elephant (*Loxodonta Africana*) Vocalizations", Spring 2003 Meetings of the Acoustical Society of America, Nashville, May 2003.
132. Heather E. Ewalt and Michael T. Johnson, "Multiple Speech Signal Enhancement using a Microphone Array", Spring 2003 Meetings of the Acoustical Society of America, Nashville, May 2003.
133. Kevin M. Indrebo, Richard J. Povinelli, and Michael T. Johnson, "A Combined Sub-band and Reconstructed Phase Space Approach to Phoneme Classification", Non-Linear Speech Processing (NOLISP) 2003, Le Croisic, France, May 2003.

134. Jinjin Ye, Michael T. Johnson, and Richard J. Povinelli, "Study of attractor variation in the reconstructed phase space of speech signals, NOLISP 2003, Le Croisic, France, May 2003.
135. Jinjin Ye, Michael T. Johnson, and Richard J. Povinelli, "Phoneme classification over reconstructed phase space using principal component analysis", NOLISP 2003, Le Croisic, France, May 2003.
136. Xiaolin Liu, Richard J. Povinelli, and Michael T. Johnson, "Vowel classification by global dynamic modeling", NOLISP 2003, Le Croisic, France, May 2003.
137. Andrew C. Lindgren, Michael T. Johnson, and Richard J. Povinelli, "Speech Recognition using Reconstructed Phase Space Features", Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2003, Hong Kong, April 2003, pp I-60 – I-63.
138. Patrick Clemins and Michael T. Johnson, "Application of Speech Recognition to African Elephant (*Loxodonta Africana*) Vocalizations", Proceedings of ICASSP 2003, Hong Kong, April 2003, pp I-484 – I-487.
139. Michael T. Johnson, Andrew C. Lindgren, Richard J. Povinelli, and Xiaolong Yuan, "Performance of Nonlinear Speech Enhancement using Phase Space Reconstruction", Proceedings of ICASSP 2003, Hong Kong, April 2003, pp I-872 – I-875.
140. Xiaolin Liu, Richard Povinelli, and Michael T. Johnson, "Detecting Determinism in Speech Phonemes", IEEE Digital Signal Processing Workshop 2002, October 2002.
141. Jinjin Ye, Richard J. Povinelli, and Michael T. Johnson, "Phoneme Classification Using naïve Bayes Classifier in Reconstructed Phase Space", IEEE Digital Signal Processing Workshop 2002, October 2002.
142. Richard J. Povinelli, Michael T. Johnson, Nabeel A.O. Demerdash, and John F. Bangura, "A Comparison of Phase Space Reconstruction and Spectral Coherence Approaches for Diagnostics of Bar and End-Ring Connector Breakage and Eccentricity Faults in Polyphase Induction Motors using Motor Design Particulars", IEEE Industry Applications Society meetings 2002, October 2002.
143. Richard J. Povinelli, Felice M. Roberts, Kristina M. Ropella, and Michael T. Johnson, "Are Nonlinear Ventricular Arrhythmia Characteristics Lost, As Signal Duration Decreases?", Computers in Cardiology 2002, September 2002.
144. Patrick Clemins and Michael T. Johnson, "Automatic Speech Recognition and Speaker Identification of Animal Vocalizations", Measuring Behavior 2002, August 2002, pp 41-46.
145. Patrick Clemins, Heather Ewalt and Michael Johnson, "Time-aligned SVD Analysis for Speaker Identification", International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2002, May 2002, p 4160.
146. Michael T. Johnson and Leah H. Jamieson, "Temporal Features for Broadcast News Segmentation", ISCA Workshop on Prosody in Speech Understanding and Recognition, October 2001.
147. Michael T. Johnson and Mary P. Harper, "Near Minimal Weighted Word Graphs for Post-processing Speech", International Workshop on Automatic speech Recognition and Understanding (ASRU), December 1999.

148. M.P. Harper, M.T. Johnson, L.H. Jamieson, S.A. Hockema, and C.M. White, "Interfacing a CDG Parser with an HMM Word Recognizer Using Word Graphs", Proceedings of the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 1999, March 1999, pp 733-736.
149. Michael T. Johnson, Leah H. Jamieson, and Mary P. Harper "Interfacing Acoustic Models with Natural Language Processing Systems", Proceedings of the International Conference on Spoken Language Processing (ICSLP) 1998, December 1998, pp 2419-2422.
150. A.M. Supranant, S.L. Hura, M.P. Harper, L.H. Jamieson, G. Long, S.M. Thede, A. Rout, T.H. Hsueh, S.A. Hockema, M.T. Johnson, J.B. Laflen, P. Srinivasa, and C.M. White, "Familiarity and Pronouncibility of Nouns and Names: The Purdue Proper Name Database", Spring 1998 Meetings of the Acoustical Society of America, June 1998.
151. Michael Johnson and Mita Desai, "An Adaptive Approach for Texture Modelling", Proceedings of the International Conference on Image Processing (ICIP) 1994, November 1994.
152. Harold Longbotham, Michael Johnson, Jack Harris, and Redouan Rouzky, "The FatBear: a nonarithmetic pico filter", Proceedings of the SPIE, Image Algebra and Morphological Image Processing IV, July 1993, pp 128-139.

Selected Invited Presentations

- Keynote Speaker, Weld Penetration Identification Based on Convolutional Neural Network, The 2019 International Workshop on Intelligentized Welding Manufacturing, University of Kentucky, Lexington, Kentucky, November 2019.
- Invited Colloquium Speaker, Acoustic to Articulatory Inversion for Computer-Aided Pronunciation Training, University of Texas at San Antonio Department of Electrical and Computer Engineering Colloquium Series, April 2019.
- Keynote Speaker, IEEE Joint Louisville-Lexington Chapters Holiday Dinner, Louisville, Kentucky, December 2018.
- Invited speaker, Shanghai Jiaotong University Colloquium Series, "Electromagnetic Articulography for Speaker Independent Acoustic-to-Articulatory Inversion, April 2015.
- Invited speaker, University of Memphis Institute of Intelligent Systems Colloquium Series, "Applying Speech Technology to Animal Vocalizations", April 2014.
- Invited speaker, NIMBioS Investigative Workshop on Analyzing Animal Vocal Sequences, "Using HMMs to model vocal structure and sequence for classification and identification", October 2013.
- Featured Keynote Speaker, GE Healthcare Technology symposium, "Innovations and Applications of Artificial Intelligence and Bayesian Modeling", August 2011.
- "Individual Identification using Hidden Markov Models for Population Monitoring and Assessment", Invited Presentation to Special Session "Topical Meeting on Signal Processing of Subtle and Complex Acoustic Signals in Animal Communication", at Acoustical Society of America Spring 2010 Meeting, Baltimore, Maryland, April 2010.
- Featured speaker at ShARE World Seminar 2008, "Innovation in Computation", Tsinghua University, Beijing China, December 2008.

- Keynote speaker for annual Tau Beta Pi banquet, “The Dr. Dolittle Project: Applying Speech Processing to Animal Vocalizations,” November 2007.
- Presentation to Beijing Institute of Technology annual colloquia, “The Dr. Dolittle Project: Applying Speech Processing to Animal Vocalizations,” Beijing, China, May 2007.
- Presentation to Tsinghua University Center for Speech Technology, “The Dr. Dolittle Project: Applying Speech Processing to Animal Vocalizations,” Beijing, China, May 2007.
- Presentation to the Chinese Academy of Sciences, “The Dr. Dolittle Project: Applying Speech Processing to Animal Vocalizations,” Institute of Acoustics, Beijing, China, May 2007.
- Michael T. Johnson, “The Dr. Dolittle Project: Applying speech technology to animal vocalizations”, Sigma Xi annual banquet, April 27, 2007.
- Michael T. Johnson, “Speech Research at Marquette – The Dr. Dolittle Project”, Marquette Trustees Committee on Academic Affairs and Planning, March 3, 2004.
- Michael T. Johnson, “Speech Recognition and Chaos Theory: A new approach to signal analysis”, Lynde Bradley Science Club, Rockwell Automation, November 11, 2003.
- Michael T. Johnson, “The Dr. Dolittle Project: Applying human speech processing algorithms to other species”, UW Milwaukee Computer Science Department Colloquium, November 7, 2003.
- Michael T. Johnson, “Talking with the Animals: Automatic Classification and Identification of Animal Vocalizations using Speech Technology”, Keynote Speaker for Physics Club of Milwaukee Annual Banquet, June 5, 2002.
- Michael T. Johnson, “Speech Processing Research at Marquette (Elephants and Cocktail Parties)”, COE National Research Council meeting, April 26, 2002.
- Michael T. Johnson, “Speech and Signal Processing Research at Marquette”, Invited speaker at Sigma Xi annual meeting, February 20, 2002.

PhD Dissertation and MS Thesis

- Michael T. Johnson, "Incorporating Prosodic Information and Language Structure into Speech Recognition Systems", PhD Dissertation, Purdue University School of Electrical and Computer Engineering, August 2000.
- Michael T. Johnson, "An Adaptive Texture Model Based on Markov Random Fields", Masters Thesis, University of Texas at San Antonio, August 1994.

Student PhD Dissertations (primary advisor)

- Soleymanpour, Mohammad, “Synthesizing Dysarthric Speech Using Multi-speaker TTS for Dysarthric Speech Recognition”, PhD Dissertation, University of Kentucky, August 2022.
- Narjes Bozorg, “Articulatory Wavenet: Deep Autoregressive Model for Acoustic-to-Articulatory Inversion”, PhD Dissertation, University of Kentucky, December 2020.

- An Ji, “Speaker Independent Acoustic-to-Articulatory Inversion”, PhD Dissertation, Marquette University, December 2014.
- Jianglin Wang, "Physiologically-motivated feature extraction methods for speaker recognition", PhD Dissertation, Marquette University, December 2013.
- Jidong Tao, “Acoustic Model Adaptation for Automatic Speech Recognition and Animal Vocalization Classification”, PhD Dissertation, Marquette University, May 2009.
- Marek Trawicki, “Distributed Multichannel Processing for Signal Enhancement”, PhD Dissertation, Marquette University, May 2009.
- Kuntoro Adi, “Hidden Markov model based animal acoustic censusing: learning from speech processing technology”, Ph.D. Dissertation, Marquette University, May 2008.
- Patrick J. Clemens, “Automatic Classification of Animal Vocalizations”, PhD Thesis, Marquette University, May 2005.

Student MS Theses (primary advisor)

- Subash Khanal, “Mispronunciation Detection and Diagnosis in Mandarin Accented English speech, Masters Thesis, University of Kentucky, May 2020.
- Jose, Neenu, “Speaker and Gender Identification using Bioacoustic Data Sets”, Masters Thesis, University of Kentucky, May 2018.
- Vonderhaar, Joe, “Speaker-specific Adaptation of Maeda Synthesis Parameters for Auditory Feedback”, Masters Thesis, Marquette University, May 2017.
- Jones, Deriq, “Development of Kinematic Templates for Automatic Pronunciation Assessment using Acoustic-to-Articulatory Inversion”, Masters Thesis, Marquette University, May 2017.
- Kelly Vonderhaar, Analysis of Interference between Electromagnetic Articulography and Electroglottograph Systems, Masters Thesis, Marquette University, August 2016.
- Xiangyu Zhou, Least-squares Mapping from Kinematic data to Acoustic Synthesis Parameters for Rehabilitative Acoustic Learning, Masters Thesis, Marquette University, May 2016.
- Andrew Kolb, Development of Software Tools and Analysis methods for the Use of Electromagnetic Articulography Data in Speech Research”, Masters Thesis, Marquette University, May 2015.
- Zac Slavens, “Generalized interpolation applied to MR image magnification and gradient nonlinearity correction”, Masters Thesis, Marquette University, May 2008.
- Xi Li, “SPEech Feature Toolbox (SPEFT) Design, and Emotional Speech Feature Extraction”, M.S. Thesis, Marquette University, August 2007.
- Anthony D. Ricke, “Automatic Frame Length, Frame Overlap, and Hidden Markov Model Topology for Automatic Speech Recognition of Animal Vocalizations”, M.S. Thesis, Marquette University, December 2006.
- Jinjin Ye, “Speech Recognition Using Time Domain Features from Phase Space Reconstructions”, M.S. Thesis, Marquette University, May 2004.
- Franck Hounkpevi, “Triphone Creation Through Rule-based Trajectory Interpolation for Continuous Speech Recognition”, M.S. Thesis, Marquette University, May 2003.

- Xiaolong Yuan, “Auditory Model-Based Bionic Wavelet Transform for Speech Enhancement”, M.S. Thesis, Marquette University, May 2003.
- Andrew Lindgren, “Speech Recognition using Features extracted from Phase Space Reconstructions”, M.S. Thesis, Marquette University, May 2003.
- Heather Ewalt, “Speech Signal Enhancement Using a Microphone Array”, M.S. Thesis, Marquette University, December 2002.

Selected News and Magazine Articles

- “Listening to the Animal Kingdom”, Discover Magazine, Marquette University, 2008, pp 14-15.
- “Decoding Calls of the Wild”, by Mark Johnson, feature article in Milwaukee Journal Sentinel, part of “Brainpower: Groundbreaking Thinkers in Wisconsin”, October 2008.
- “Say What?”, cover article on Dr. Dolittle Project, Weekly Reader Science, Spring 2007.
- “Animal Talk”, by Elena Cabral, Scholastic News Edition 5/6, December 18, 2006.
- “Animal “Speech” Project Aims to Decode Critter Communication”, by Maryann Mott, National Geographic Online, <http://news.nationalgeographic.com/news/2006/09/060926-dolittle-project.html>, September 2006.
- “Parsing the Puffin’s Patois”, article on Dr. Dolittle Project by Rachel Metz, Wired News online, July 2006.
- “Quacking the Code”, by Katie Pelech, article in Milwaukee Magazine, March 2006, p. 30.
- “What Jumbo tells Dumbo” by Alex Antunes, feature article in Computers in Science and Engineering (CISE) magazine, published by IEEE Computer Society, September-October 2005, pp 3-6.
- “Discoveries with Disney”, article about collaborative project with Disney’s Animal Kingdom published in MU Magazine, Fall 2003, p. 17.

GRANTS

Major Grants Funded

1. Michael T. Johnson (PI), Aaron Cramer (co-PI), Dan Ionel (co-PI), and Janet Lumpp (co-PI), GAANN Fellowship Program in Electrical and Computer Engineering at the University of Kentucky, Department of Education, 10/1/2019 – 9/30/2022. Total Budget \$600,702 over 3 years. Supplement of \$51,000 in 2020 for 1 additional fellow.
2. Jeffrey Berry (PI) and Michael T. Johnson (co-PI), “Speech Sensorimotor Control in Non-progressive Dysarthria”, National Institutes of Health R-15, 7/1/2018 – 6/30/21. Total Budget \$446,956 over 3 years (\$166,428 to University of Kentucky as subcontract to Marquette University).
3. Michael T. Johnson (PI) and Jeffrey Berry (co-PI), “RI: Small: Speaker Independent Acoustic-Articulator Inversion for Pronunciation Assessment”, National Science Foundation CISE Directorate, 2013. Total Budget \$449,643 over 3 years.
4. Michael T. Johnson (PI) and Jeffrey Berry (co-PI), “EAGER: Acoustic-Articulator Modeling for Pronunciation Analysis”, National Science Foundation CISE Directorate, 2011. Total Budget: \$149,896.00 over 1 year.
5. Michael T. Johnson (PI), Elizabeth Mugganthalder (Fauna Communications Research Institute), Peter Scheifele (National Undersea Research Center, University of Connecticut), Mike Darre (Animal Sciences, University of Connecticut), Anne Savage (Disney Animal Kingdom), “The Dr. Dolittle Project: A Framework for Classification and Understanding of Animal Vocalizations”, NSF IT-R program under CISE Directorate, 2003. Total Budget: \$1,200,000 over 4 years.
6. Michael T. Johnson (PI) and Richard J. Povinelli, “Integration of Stochastic and Dynamical Methods for Speech Technology”, NSF IT-R program under CISE directorate, 2001. Total Budget: \$360,000 over 3 years.
7. Richard J. Povinelli (PI), Nabeel A. O. Demerdash, Edwin Yaz, and Michael T. Johnson, “A Novel Approach to Fault Modeling, Diagnostics, and Prediction in Motor Drive Systems”, NSF Division of Controls, Networks, and Computational Intelligence, 2003. Total Budget: \$340,000 over 3 years.

Additional Grants Funded

1. Peter Scheifele (PI), Leesa Scheifele (co-PI), and Michael T. Johnson (co-PI), Measuring Hearing Acuity and Vocalization Signals of the Asian Small-Clawed Otter (*Amblonyx cinereus*) in Response to Anthropogenic Noise, Indianapolis Zoo, March 2017, Funded \$14,300.
2. Kevin McGowan (PI) and Michael T. Johnson (co-PI), Apparatus and Procedure for Ultrasound Investigation of Velopharyngeal port aperture, University of Kentucky Igniting Research Collaborations project, Funded \$27,726 in 2018-2019.
3. Jeffrey Berry (PI) and Michael T. Johnson (co-PI), “Speech Rehabilitation Using a Virtual Vocal Tract”, Marquette Regular Research Grant (RRG). Funded \$6,000 in 2014-15.
4. Michael T. Johnson (PI) and Jeffrey Berry (co-PI), Research Experiences for Undergraduates (REU) supplement to “RI: Small: Speaker Independent Acoustic-Articulator Inversion for Pronunciation Assessment”, NSF REU program, CISE directorate. Funded \$6,000 in 2013-2014, \$6000 in 2014-2015, \$6000 in 2015-2016.
5. Michael T. Johnson (PI) and Jeffrey Berry (co-PI), OISE International supplement to “EAGER: Acoustic-Articulator Modeling for Pronunciation Analysis”, co-funded by CISE and OISE directorates. Funded \$19,840 in 2012-2013.

6. Michael T. Johnson (PI) and Jeffrey Berry (co-PI), Research Experiences for Undergraduates (REU) supplement to “EAGER: Acoustic-Articulator Modeling for Pronunciation Analysis”, NSF REU program, CISE directorate. Funded \$8,000 in 2011-2012.
7. Richard J. Povinelli (PI) and Michael T. Johnson (PI), “An examination of phoneme confusability in spoken English”, Marquette Regular Research Grant (RRG). Funded \$15,000 in 2011-2012.
8. Michael T. Johnson (PI) et. al., Research Experiences for Undergraduates (REU) supplement to “The Dr. Dolittle Project”, NSF REU program, CISE directorate. Funded at \$6,000 in 2003-04, \$15,000 in 2004-2005, \$12,000 in 2005-2006, \$15,000 in 2006-2007 Total Budget: \$48,000.
9. Michael T. Johnson (PI) and Richard J. Povinelli, REU supplement to “Integration of Stochastic and Dynamical Methods for Speech Technology”, NSF REU program, CISE directorate. Funded at \$15,000 in 2001-02, 2002-03, and 2003-04, Total Budget: \$45,000.
10. Richard J. Povinelli (PI), Nabeel A. O. Demerdash, Edwin Yaz, and Michael T. Johnson, REU supplement to “A Novel Approach to Fault Modeling, Diagnostics, and Prediction in Motor Drive Systems”, NSF REU Program, CNCI Division. Funded at \$6,000 in 2003-04 and 2004-05. Total Budget to-date: \$6,000.

PROFESSIONAL ACTIVITIES

Technical Committees and Publication editing

- Area Editor, Speech Enhancement, Speech Communication Journal, 2014-present
- Member, IEEE Signal Processing Society Speech and Language Technical Committee (SLTC), 2014-2017

Selected Publication Review Participation

- Speech Communication Journal
- International Conference on Acoustics, Speech and Signal Processing
- Interspeech IEEE Automatic Speech and Recognition Workshop
- Workshop on Spoken Language Technology
- IEEE Transactions on Speech and Audio Processing
- Journal of the Acoustical Society of America
- IEEE Sensors Journal
- IEEE Transactions on VLSI
- IEEE Transactions on Signal Processing
- IEEE Signal Processing Letters
- IEEE Transactions on Neural Systems
- Journal of Applications of Signal Processing
- Animal Behavior Journal
- EURASIP Journal on Audio, Speech and Music Processing
- International Journal on Adaptive Control and Signal Processing
- Biosystems Engineering Journal
- Acoustics Research Letters Online
- International Conference on Decision and Control
- Midwest Symposium on Circuits and Systems
- Iasted Circuits and Systems Conference
- Workshop on Signal Processing and Applications
- American Controls Conference
- Journal of Zhejiang University Science C (Computers & Electronics)
- Journal of Medical Engineering & Physics
- Open Signal Processing Journal
- IEEE Spoken Language Technology workshop
- International Conference on Signal Processing (ICSP)
- Journal of Computers
- International Journal of Electronics and Communications
- Franklin Institute journal

Selected Funding Review Participation

- National Science Foundation
Have reviewed for CCLI, CAREER, CONICYT, and ITR programs, as well as various regular programs in the CISE and Biological Sciences directorate
- International: Austrian Science Fund, Icelandic Research Fund
- Medical College of Wisconsin Clinical and Translational Science Institute (CTSI)

Professional and honorary society memberships

- ECE Department Heads Association (ECEDHA), 2016-present
ECEDHA Awards Committee, 2018-present
- Southeast ECEDHA (SECEDHA), 2016-present
Secretary 2019-2020, Vice President 2020-2022, President 2022-2023
- Southeastern Center for EE Education (SCEEE) Board of Directors, 2020-2023
- Senior Member IEEE, membership 1994-present
- IEEE Signal Processing Society, 1998-present
- Acoustical Society of America (ASA), 2001-present
- Eta Kappa Nu, 1996-present
- Upsilon Pi Epsilon, 2001-present
- Senior Member Sigma Xi, 2001-2012
- Association of Computer Machinery (ACM), 1998-2010
- IEEE Computer Society, 1998-2010
- IEEE Circuits and Systems Society, 2001-2009
- Association of Computational Linguistics (ACL), 1998-2002
- International Speech Communication Association (ISCA), 2001-2006

Other professional activities and awards

- Selected for SEC Academic Leadership Development Program, 2019-2020
- Featured in “Groundbreaking Thinkers in Wisconsin” article series 2007
- Researcher of the Year, Marquette College of Engineering, 2006-2007
- Young Engineer of the Year, Engineers and Scientists of Milwaukee (ESM), 2006
- Registered Professional Engineer, State of Wisconsin

TEACHING AND EDUCATIONAL ACTIVITIES

Courses Taught

Undergraduate

EECE Freshman Seminar
Computer Hardware
Digital Logic Design
Linear Systems Analysis
Embedded Systems Design
Circuits 1
Signals and Systems Lab
Introduction to Engineering
Understanding Leadership

Graduate and Joint Undergraduate/Graduate

Digital Signal Processing
Speech Processing
Optimal/Adaptive Signal Processing
Pattern Recognition
Analysis of Algorithms
Information Theory
Professional Research Writing

Teaching awards

Marquette Eta Kappa Nu (HKN) honor society EECE Teacher of the Year, 2014, 2005

Selected Teaching, Education, and Pedagogy activities

UK Global Engagement Academy, Spring 2022
Center for Enhancement of Learning and Teaching (CELT) workshop, Spring 2021
UK Counseling Center, ECE faculty meeting guest on Student Mental Health, Fall 2021
UK Unconscious Bias Seminar, Fall 2019
UK Experienced Leadership Academy, Fall 2018
UK CoE Teaching Community of Practice, Fall 2017
UK Chairs Academy, Fall 2016
Chair, CoE Teaching Best Practices Committee, Spring 2017
Coordinator for Marquette EECE Freshman Seminar, 2011-2015
Participant in KEEN/Lafferty Teaching Development Workshop, Fall 2015, Spring 2016
Mentored numerous Research Experience for Undergraduates (REU) through NSF grants
International Teaching Experience, Tsinghua University, 2008-2009, 2011, 2013, 2014-2015
Signals and Systems Concept Inventory site coordinator for NSF study, 2003-2006
Developed and gave numerous DSP/Audio Workshops for middle and high school outreach
Participated with IEEE Signal Processing Education Technical Committee activities
Developed introductory DSP module for General Engineering freshman course
Participated in MU Assessment seminar by Barbara Wolvoord, May 2005
Participated in NSF Engineering Education Scholars (EES) program for new faculty
Participated in Prentice Hall Symposium on Education, Chicago, IL, 2001
Participated in NSF/ASEE Visiting Scholars Teaching Workshop, Fall 2000 and Spring 2001
Member, Advisory committee MU Center on Teaching and Learning (2005-2010)

Student Mentoring Activities (2000-present)

MS and PhD Committees
Undergraduate Research Projects
Senior Design Projects
Individual Student Mentoring

ACADEMIC SERVICE

Academic Service at the University of Kentucky

University:

- Steering Committee for Wildcat Foundations first-year experience program (2017-2019)
- Represented UK in SEC Academic Leadership Development Program (ALDP), 2019-2020

College:

- Chair, Scholarship Processes Committee, Engineering (2023-2024)
- Member, Funkhouser Renovation Planning Committee (2023-2024)
- Advocate, Advocates and Allies Working Group (2023-2024)
- Chair, BME Department Chair Search Committee (2022-2023)
- Chair, Engineering Technology Faculty Committee (2020 – present)
ET Chair search, LST Faculty search (2020-2021), CPT Faculty search (2021-2022)
- CoE Funkhouser (2019, 2021, 2023) and Grehan (2018-19) Building Committees
- CoE Recruitment Advisory Committee (2019-present)
- CoE Search Committee for Senior Director of Philanthropy (2018)
- Director, CoE Scholars in Engineering Leadership Program (2017-present)
- CoE Search Committee for Director of Marketing and Communications (2017)
- Chair, CoE Teaching Best Practices committee (Spring 2017)
- CoE Chair's Committee (2016-present, Chair 2018-2019)

Departmental:

- ECE Department Chair (2016-2023)

Academic Service at Marquette University

University:

- University Board of Graduate Studies (2009-2016; chair, 2012-2016)
- University Search Committee, Vice Provost for Enrollment Management (2015-2016)
- University Steering committee on academic integrity (2012-2014)
- University Subcommittee on academic integrity (2011)
- COF, Subcommittee on Nominations and Elections (chair, 2004-2007).
- Faculty mentor for Dissertation Boot Camps (2008, 2010)
- Advisory Committee for Center on Teaching and Learning (2005-2010)

College and Departmental:

- COE Research Activity Committee (2011)
- COE Dean of Engineering Search Committees (2003, 2009)
- COE Freshman Programs Committee (2006-2007)
- COE Discovery Towers Building Committee (2004-2005)
- BME Biocomputer Program Advisory Board (2001-2016)
- EECE Graduate Committee (2000-2003, 2005-2016)
- EECE Undergraduate Committee (2003-2008, 2012-2015)
- EECE Director of Graduate Studies (2009-2012)
- EECE Department Grad Program Assessment Coordinator (2009-2012)
- Numerous EECE faculty search and other subcommittees

Selected public service and outreach activities

- UK CoE E-day, every spring annually since 2016
- UK CoE and ECE prospective student tours, several times per month since 2016
- Guest speaker, Triangle UK Leadership Academy students, July 2023
- Keynote Speaker, International Workshop on Intelligentized Welding Manufacturing, November 2019
- Keynote Speaker, IEEE Joint Lexington/Louisville Section Meeting, December 2018
- UK ECE Middle School STEM program, November 2018
- UK ECE STEM Outreach camp for girls, June 2016, June 2017
- Guest speaker, Marquette Girls Who Code outreach program, March 2016.
- Coordinator for Marquette EECE open house activities, Fall 2015.
- Faculty advisor for Marquette's HKN chapter and Sigma Phi Delta engineering fraternity
- Guest speaker, MOTIF English corner, Beijing, China. "Diversity", September 2014, "Humor", March 2015, "Languages of the World", June 2015.
- Senior Visiting Scholar, Tsinghua University, China
2014-15, 2008-09, Summer 2011, Summer 2013
- COE iPalm2 outreach program to high school students, August 2011, 2012, 2013
- Represented Marquette COE on visits to Harbin Institute of Technology, China, 2011
- Panelist for Marquette University Advancement workshop, 2011
- Marquette ORSP Brown Bag panelist "One thing led to another" research series, 2010
- Spoke to students at Rosenwald Dunbar elementary school in Nicholasville, KY, about the Dolittle project and careers in engineering, 2008
- Spoke to students at Cardinal Valley elementary school in Lexington, KY, about the Dolittle project and careers in engineering, 2006
- Participated in Marquette University High School Junior Career Day, 2005
- Spoke to "Future Problem Solvers" gifted program at Templeton Middle School about Artificial Intelligence, 2005
- COE representative for "We are Marquette" World Café Conversation, 2004
- EECE representative and presenter for GEMS Advanced Education Fair, 2004
- Spoke at MSOE about graduate school opportunities in EECE, 2002-2003
- Presented at Lynde Bradley Science Club, Rockwell Automation, 2003
- Presented at UW Milwaukee Computer Science Department Colloquium, 2003
- Marquette ORSP Brown Bag panelist on Pre-proposal Contact with Sponsors, 2003
- Keynote on Global Research, Marquette COE National Advisory Council, 2002
- Keynote Speaker at Physics Club of Milwaukee annual banquet, 2002
- Spoke at UWM Elementary Education class on careers in science and technology, 2002
- Participated in Nathan Hale H.S. Career Day regarding careers in engineering, 2002
- Active participant in Marquette EECE and COE open houses, 2000-2015
- Blackjack dealer for Marquette engineering student council casino nights, 2000-15
- Regular participant in Marquette parent lunch program, 2000-2015
- Member, Eastbrook Church & worship choir, 2001-2016
- Numerous student and student group activities, both UK and Marquette, 2000-present
- Numerous research talks, colloquia, and seminars, 2000-present
- Numerous other volunteer and community activities, 2000-present

LITIGATION EXPERIENCE

US District Court

Snap-On Technologies, Inc. vs. Hunter Engr. Company, 98-C-369, 98-C-837, 00-C-929, 00-C-1511 (2002)

Jurisdiction: USDC Eastern District of Texas
Client: Hunger Engineering Company (Defendant)
Law Firm: Kirkland and Ellis LLP
Activities: Declaration, Deposition

EVS Codec Tech LLC and Saint Lawrence Comm. LLC vs. ZTE and ZTE USA Inc., 3:19-cv-00385 (2019)

Jurisdiction: USDC Northern District of Texas
Client: ZTE Inc. (Defendant)
Law Firm: K&L Gates LLP
Activities: Source Code Review, Technical consulting services

Nuance Communications, Inc. vs. MModal LLC., 1:17-cv-01484-MN (2019)

Jurisdiction: USDC District of Delaware
Client: MModal LLC (Defendant)
Law Firm: Duane Morris, LLP and Latham & Watkins, LLP
Activities: Reports (3), Source Code Review, Deposition

VoiceAge EVS LLC vs. Apple Inc, 1:20-cv-01061-CFC (2020)

Jurisdiction: USDC District of Delaware
Client: Apple Inc. (Defendant)
Law Firm: WilmerHale LLP
Activities: Source Code Review, Technical consulting services

Freshhub vs. Amazon Inc., 1:19-cv-00885-ADA (2020)

Jurisdiction: USDC Western District of Texas
Client: Amazon (Defendant)
Law Firm: Fenwick and West LLP
Activities: Reports (3), Source Code Review, Depositions (2), Trial Testimony

Vocalife LLC vs. Sonos, Inc., 2:21-cv-00129 (2021)

Jurisdiction: USDC Eastern District of Texas
Client: Sonos (Defendant)
Law Firm: Lee, Sullivan, Shea, and Smith LLP
Activities: Declaration

VB Assets, LLC vs. Amazon.com, Inc., 1:19-cv-01410-MN (2021)

Jurisdiction: USDC District of Delaware
Type of Litigation: Patent Infringement
Client: Amazon (Defendant)
Law Firm: Fenwick and West LLP
Activities: Reports (3), Depositions (2), Trial Testimony

Dialect, LLC v. Amazon.com et al., No. 1:23-cv-581 (2024)

Jurisdiction: USDC Eastern District of Virginia
Type of Litigation: Patent Infringement
Client: Amazon (Defendant)
Law Firm: Fenwick and West LLP
Activities: Technical Consulting (currently active)

Patent Trial and Appeal Board (PTAB) Inter Partes Review (IPR)

LG Electronics vs. Saint Lawrence Communication, IPR2015-01874, IPR2015-01875 (2015)

Client: LG Electronics (Petitioner)

Law Firm: Mayer Brown LLP

Activities: Declarations (2)

ZTE Corp. vs. Saint Lawrence Communications, IPR2016-00704, IPR2016-00705 (2016)

Client: ZTE Inc. (Petitioner)

Law Firm: Finnegan, Henderson, Farabow, Garrett & Dunner LLP

Activities: Declarations (2), Deposition

Google LLC vs. Koninklijke Philips, IPR2017-00437 (2017)

Client: Koninklijke Philips (Patent Owner)

Law Firm: Fitzpatrick, Cella, Harper, & Scinto

Activities: Declaration, Deposition

Google LLP vs. Sonos, Inc., IPR2021-00475 (2022)

Client: Sonos (Patent Owner)

Law Firm: Ackerman LLP and Lee, Sullivan, Shea, and Smith LLP

Activities: Declaration

Sonos, Inc. vs. Google LLP, IPR2022-01592, 2022-01594, 2023-0118, 2023-0119 (2022)

Client: Sonos (Petitioner)

Law Firm: Lee, Sullivan, Shea, and Smith LLP

Activities: Declarations (4) (currently active, all 4 IPRs instituted)

Sonos, Inc. vs. Google LLP, IPR2023-01288, IPR2023-01284 (2023)

Client: Sonos (Petitioner)

Law Firm: Lee, Sullivan, Shea, and Smith LLP

Activities: IPR Declarations (2)

Jawbone Innovations, LLC vs. Meta Platforms, Inc., IPR2024-00349, IPR2024-00350, (2024)

Jurisdiction: USDC Western District of Texas

Client: Meta Platforms (Petitioner)

Law Firm: Allen & Overy LLP

Activities: IPR Declarations (2)

International Trade Commission (ITC)

Google LLP vs. Sonos, Inc. 337-TA-1329, 337-TA-1330 (2022)

Client: Sonos (Respondent)

Law Firm: Lee, Sullivan, Shea, and Smith LLP

Activities: Technical Consulting, 4 reports, 2 depositions, ITC hearing testimony

Other Jurisdictions

Google LLP vs. Sonos, Inc. Docket T-952-20 (2021)

Jurisdiction: Canadian Federal Court

Type of Litigation: Patent Infringement

Client: Sonos (Defendant)

Law Firm: ROBIC, LLP and Lee, Sullivan, Shea, and Smith LLP

Activities: Reports (2), Source Code Review, Trial Testimony

APPENDIX B

Description
U.S. Patent No. 8,805,692 (“’692”)
File History of U.S. Application No. 14/148,752 (“’692FH”)
U.S. Patent Application Publication No. 2006/0132382 (“Jannard”)
U.S. Patent 7,529,677 B1 (“Wittenberg”)
Wittenberg Patent Application (No. 11/038,426) As-Filed
U.S. Patent Application Publication No. 2006/0020960 (“Relan”)
U.S. Patent Application Publication No. 2003/0220835 (“Barnes”)
U.S. Patent Application Publication No. 2003/0130852 (“Tanaka”)
U.S. Patent Application Publication No. 2010/0311390 (“Black”).
U.S. Patent Application Publication No. 2007/0165875 (“Rezvani”)
U.S. Patent Application Publication No. 2006/0136378 (“Martin”)
Australian Patent Application No. 2005201979 A1
Computer Dictionary, 2 nd Edition, 1994, excerpt page 365.
U.S. Patent No. 5,251,263 (“Andrea”)
Provisional Patent Application No. 60/806,769
Spoken language Processing, Huang, 2001, Section 10.2, “Acoustical Transducers,” page 482