

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.,
Petitioner,

v.

EIGHT KHZ, LLC,
Patent Owner.

IPR2023-01004
Patent 9,282,196 B1

Before JUSTIN T. ARBES, LYNNE H. BROWNE, and
SCOTT RAEVSKY, *Administrative Patent Judges*.

ARBES, *Administrative Patent Judge*.

DECISION
Denying Institution of *Inter Partes* Review
35 U.S.C. § 314

I. INTRODUCTION

A. Background and Summary

Petitioner Meta Platforms, Inc. filed a Petition (Paper 1, “Pet.”) requesting *inter partes* review of claims 1, 4–6, 8–11, 14, 15, and 17–19 of U.S. Patent No. 9,282,196 B1 (Ex. 1001, “the ’196 patent”) pursuant to 35 U.S.C. § 311(a). Patent Owner Eight kHz, LLC filed a Preliminary Response (Paper 6, “Prelim. Resp.”) pursuant to 35 U.S.C. § 313. Pursuant

to our authorization (Ex. 1045), Petitioner filed a Reply (Paper 7, “Reply”) and Patent Owner filed a Sur-Reply (Paper 8, “Sur-Reply”).

Pursuant to 35 U.S.C. § 314(a), the Director may not authorize an *inter partes* review unless the information in the petition and preliminary response “shows that there is a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition.” *See* 37 C.F.R. § 42.4(a) (“The Board institutes the trial on behalf of the Director.”). For the reasons that follow, we do not institute an *inter partes* review.

B. Related Matters

The parties indicate that the ’196 patent is the subject of *Eight kHz, LLC v. Meta Platforms, Inc.*, No. 6:22-cv-00575 (W.D. Tex.) (“the district court case”). *See* Pet. 3; Paper 4, 2. Petitioner filed petitions challenging patents related to the ’196 patent in Cases IPR2023-01003 (“the -1003 case”) and IPR2023-01005, and petitions challenging unrelated patents asserted in the district court case in Cases IPR2023-01019, IPR2023-01020, IPR2023-01021, IPR2023-01022, IPR2023-01023, and IPR2023-01024.

C. The ’196 Patent

The ’196 patent states that electronic devices, such as telephones, “typically provide monophonic or stereophonic sound,” which has “good speech intelligibility” but “is not comparable in quality to sound that the listeners would hear if they were communicating face-to-face with each other.” Ex. 1001, col. 1, ll. 8–16. The ’196 patent discloses an improved “computer system in which two electronic devices communicate with each other during a telephone call between two people” and “[t]he computer

system positions a voice of one person at a sound localization point that is proximate to the other person.” *Id.* at col. 2, ll. 42–46.

In order to determine a location of sound, humans process and compare monaural cues from each ear. This comparison reveals difference cues or binaural cues that enable sound localization. These cues include interaural time differences (ITDs), interaural level differences (ILDs), and head-related transfer functions (HRTFs). Cues thus result from changes to the sound wave from an interaction with the human anatomy. Humans process these changes as impulse responses or head-related impulse responses (HRIRs). Once the HRIRs for an individual are known, the associated or transformed HRTFs can be calculated. The HRIRs and transformed HRTFs enable sound to be convolved such that a location of the sound or sound localization point can be changed. A listener hears the convolved sound as if it had originated from the sound localization point.

Id. at col. 2, ll. 53–67.

The ’196 patent describes, with reference to Figure 1, a method of “provid[ing] sound that is localized at a sound localization point.” *Id.* at col. 3, ll. 1–2. Step 100 is to “capture sound that will be provided to a listener,” where the “sound can be provided in real-time to a listener, processed and delayed, or stored in memory.” *Id.* at col. 3, ll. 3–9. Step 110 is to “designate a sound localization point at a location in empty space that is away from and proximate to a location of the listener such that the sound localization point simulates an origin of the sound that the listener hears.” *Id.* at col. 3, ll. 33–36. The sound localization point (SLP) can be a three-dimensional or two-dimensional position. *Id.* at col. 3, ll. 46–54. Step 120 is to “adjust the sound so the sound originates from the sound localization point at the location in empty space that is away from and proximate to the location of the listener.” *Id.* at col. 4, ll. 16–19. For example, the sound can be “processed to alter, add, or generate head-related

transfer functions (HRTFs), interaural time differences, and/or interaural level differences to correspond to a fixed or variable distance from the listener, a coordinate location or a GPS location of the sound localization point.” *Id.* at col. 4, ll. 24–29. Lastly, step 130 is to “provide the sound to the listener such that the listener localizes the origin of the sound at the sound localization point in empty space.” *Id.* at col. 4, ll. 30–32. The SLP can be “an empty space” or an occupied space where a physical or virtual object or character exists. *Id.* at col. 4, ll. 33–59.

Similarly, the ’196 patent describes, with reference to Figure 9, a method of “adust[ing] HRTFs of an individual during an electronic call.” *Id.* at col. 21, ll. 7–8. Step 900 is to “execute an electronic call between a first individual and a second individual such that an origin of a voice of the second individual appears to the first individual to originate from a sound localization point that simulates a location in empty space of the second individual.” *Id.* at col. 21, ll. 9–13. Step 910 is to “monitor, during the electronic call, movements of a head location and/or head orientation of the first individual in response to the first individual localizing sounds,” for example, using an electronic device that determines “when the [first individual’s] head turns left or right, turns up or down, or rotates back and forth.” *Id.* at col. 21, ll. 24–32. Lastly, step 920 is to “adjust, during the electronic call, stored [HRTFs] of the first individual to more accurately simulate real HRTFs of the first individual in response to the movements of the head location and/or head orientation of the first individual.” *Id.* at col. 21, ll. 43–47. The HRTFs can be adjusted in real time. *Id.* at col. 21, ll. 51–53.

Figure 17 of the '196 patent is reproduced below.

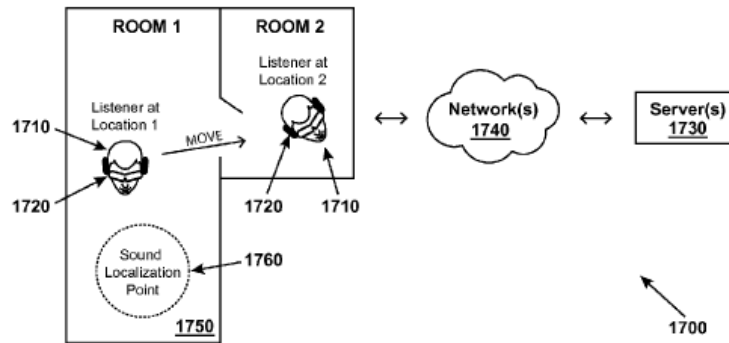


Figure 17

Figure 17 depicts listener 1710 wearing wearable electronic device (WED) 1720 that plays sound, displays corresponding images, and tracks the location and head orientation of the listener. *Id.* at col. 30, ll. 3–6, 14–18, 37–43. Sound from WED 1720 initially localizes at SLP 1760 in empty space of room 1. *Id.* at col. 30, ll. 6–10. When the listener moves, SLP 1760 “remains fixed at the location in room 1” and the sound “automatically adjusts to compensate for changes to the head orientation and the location of the listener as the listener moves about rooms 1 and 2.” *Id.* at col. 30, ll. 10–35. “For example, amplitude and localization cues (e.g., ITD and ILD) change in real-time as the listener 1710 moves from room 1 to room 2.” *Id.* at col. 30, ll. 35–37.

Figure 18 of the '196 patent is reproduced below.

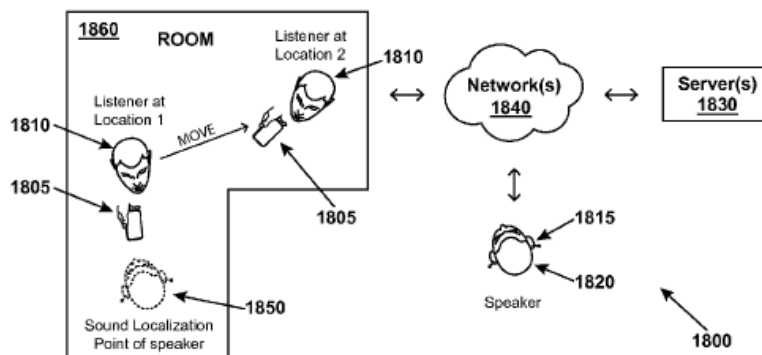


Figure 18

Figure 18 depicts a similar arrangement involving listener 1810 conducting an electronic call with handheld portable electronic device (HPED) 1805 that captures and emits binaural sound and tracks the location and head orientation of the listener. *Id.* at col. 30, l. 60–col. 31, l. 4. SLP 1850 “has a size, a shape, and an orientation of a head that emulates a size, a shape, and an orientation of a head of the speaker 1820” such that “[w]hen the speaker 1820 moves his head (such as changing location, yaw, pitch, or roll), then a head of the sound localization point 1850 simultaneously changes to emulate or copy this movement” and “[a] direction or source of the sound from the sound localization point emulates sound from the mouth of the speaker 1820.” *Id.* at col. 31, ll. 5–12. Likewise,

[w]hen the listener 1810 moves his head with respect to the sound localization point 1850 (such as changing location, yaw, pitch, or roll), then adjustments are simultaneously made to the sound of the voice of the speaker to compensate for these movements. These adjustments change the sound so the sound that the listener hears emulates the sound that the listener would hear if the speaker were physically present at the sound localization point 1850. For instance, the listener 1810 continues to hear the voice of the speaker 1820 at the sound localization point as the listener moves from location 1 in the room 1860 to location 2 in the room 1860.

Id. at col. 31, ll. 21–31.

Finally, the ’196 patent discloses an embodiment in which a user speaks with others on a conference call and the system localizes the voice of each user at different SLPs around a table (e.g., at empty spaces or chairs at the table). *Id.* at col. 34, l. 58–col. 35, l. 44, Fig. 23. “Adjustments are made to the sound for each of the[] users since they are located at different locations around the table with respect to the sound localization points” (e.g., adjustments “based on movements to head orientation and location for

each user”). *Id.* at col. 35, ll. 49–63. The ’196 patent also discloses an example in which a user is speaking to four others that “are not people but computer programs with interactive audio input/output interfaces.” *Id.* at col. 35, l. 64–col. 36, l. 1. The user “can hear each of his programs running in a different spot on his large desk” and “can move the SLPs of a window/process using, for example, a voice command, a mouse, a head or other body gesture, or other method.” *Id.* at col. 36, ll. 12–17.

D. Illustrative Claim

Challenged claims 1, 8, and 15 of the ’196 patent are independent. Claims 4–6 depend from claim 1, claims 9–11 and 14 depend from claim 8, and claims 17–19 depend from claim 15. Claim 1 recites (with letter designations used in the Petition to refer to the various limitations):

1. [pre] A method executed by one or more electronic devices in an electronic system that move a sound localization point (SLP) of a voice of a computer program during a voice exchange between the computer program and a user, the method comprising:

[a] providing, through an electronic earphone being located at a head of the user and being in communication with a handheld portable electronic device (HPED), binaural sound to the user during the voice exchange between the computer program and the user such that an origin of the voice of the computer program occurs at the SLP at a first location in empty space that is away from and proximate to the user;

[b] receiving, from the user and at the HPED, a command to move the SLP of the voice of the computer program from the first location in empty space to a second location in empty space that is proximate to the first location and away from and proximate to the user;

[c] moving, based on the command and during the voice exchange between the computer program and the user, the SLP

of the voice of the computer program from the first location in empty space to the second location in empty space that is proximate to the first location and away from and proximate to the user; and

[d] providing, through the electronic earphone being located at the head of the user and being in communication with the HPED, the binaural sound to the user during the voice exchange between the computer program and the user such that the origin of the voice of the computer program occurs at the SLP at the second location in empty space that is proximate to the first location and away from and proximate to the user.

E. Evidence

Petitioner relies on the following prior art:

U.S. Patent No. 8,406,439 B1, issued Mar. 26, 2013 (Ex. 1005, “Bedingfield”);

U.S. Patent No. 9,131,305 B2, filed Jan. 17, 2013, issued Sept. 8, 2015 (Ex. 1006, “Li”);

U.S. Patent Application Publication No. 2003/0059070 A1, published Mar. 27, 2003 (Ex. 1008, “Ballas”);

U.S. Patent Application Publication No. 2008/0273708 A1, published Nov. 6, 2008 (Ex. 1017, “Sandgren”); and

U.S. Patent Application Publication No. 2009/0060235 A1, published Mar. 5, 2009 (Ex. 1016, “Woo”).

F. Prior Art and Asserted Grounds

Petitioner asserts that claims 1, 4–6, 8–11, 14, 15, and 17–19 of the ’196 patent are unpatentable on the following grounds:

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
1, 8, 15, 17	103	Li
4, 6, 9, 10, 14, 18	103	Li, Ballas

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
5	103	Li, Ballas, Sandgren
19	103	Li, Sandgren
11	103	Li, Woo
1, 8, 15, 17	103	Bedingfield
4, 6, 9, 10, 14, 18	103	Bedingfield, Ballas
5	103	Bedingfield, Ballas, Sandgren
19	103	Bedingfield, Sandgren
11	103	Bedingfield, Woo

II. ANALYSIS

A. Legal Standards

A claim is unpatentable for obviousness if “the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains.” 35 U.S.C. § 103; *see KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations, including “the scope and content of the prior art”; “differences between the prior art and the claims at issue”; and “the level of ordinary skill in the pertinent art.” *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966). Additionally, secondary considerations, such as “commercial success, long felt but unsolved needs, failure of others, etc., might be utilized to give light to the circumstances surrounding the origin of the subject matter sought to be patented. As indicia of obviousness or nonobviousness, these inquiries may have

relevancy.” *Id.* When conducting an obviousness analysis, we consider a prior art reference “not only for what it expressly teaches, but also for what it fairly suggests.” *Bradium Techs. LLC v. Iancu*, 923 F.3d 1032, 1049 (Fed. Cir. 2019).

A patent claim “is not proved obvious merely by demonstrating that each of its elements was, independently, known in the prior art.” *KSR*, 550 U.S. at 418. An obviousness determination requires finding “both ‘that a skilled artisan would have been motivated to combine the teachings of the prior art references to achieve the claimed invention, and that the skilled artisan would have had a reasonable expectation of success in doing so.’” *Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359, 1367–68 (Fed. Cir. 2016); *see KSR*, 550 U.S. at 418 (for an obviousness analysis, “it can be important to identify a reason that would have prompted a person of ordinary skill in the relevant field to combine the elements in the way the claimed new invention does”). Also, “[t]hough less common, in appropriate circumstances, a patent can be obvious in light of a single prior art reference if it would have been obvious to modify that reference to arrive at the patented invention.” *Arendi S.A.R.L. v. Apple Inc.*, 832 F.3d 1355, 1361 (Fed. Cir. 2016).

“Although the *KSR* test is flexible, the Board ‘must still be careful not to allow hindsight reconstruction of references . . . without any explanation as to *how* or *why* the references would be combined to produce the claimed invention.’” *TriVascular, Inc. v. Samuels*, 812 F.3d 1056, 1066 (Fed. Cir. 2016). Further, an assertion of obviousness “cannot be sustained by mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness.” *KSR*, 550 U.S. at 418 (quoting *In re Kahn*, 441 F.3d 977, 988

(Fed. Cir. 2006)); *accord In re NuVasive, Inc.*, 842 F.3d 1376, 1383 (Fed. Cir. 2016) (stating that “conclusory statements” amount to an “insufficient articulation[] of motivation to combine”; “instead, the finding must be supported by a ‘reasoned explanation’”); *In re Magnum Oil Tools Int’l, Ltd.*, 829 F.3d 1364, 1380 (Fed. Cir. 2016) (“To satisfy its burden of proving obviousness, a petitioner cannot employ mere conclusory statements. The petitioner must instead articulate specific reasoning, based on evidence of record, to support the legal conclusion of obviousness.”).

B. Level of Ordinary Skill in the Art

Petitioner argues that at the time of the ’196 patent (June 2014), a person of ordinary skill in the art would have had “at least a bachelor’s degree in computer science or engineering with a focus on audio signal processing and sound localization, with at least two years of relevant experience,” where “[a]dditional experience could substitute for less education, and additional education could likewise substitute for less experience.” Pet. 14; *see* Ex. 1003 ¶¶ 30–35. Patent Owner argues that a person of ordinary skill in the art instead would have had “several years of academic and/or professional experience in a technical field relevant to digital audio signal processing, such as electrical engineering, computer engineering, or computer science; and a familiarity with relevant literature and concepts concerning virtual 3D audio processing, which were widely available at the time of disclosure.” Prelim. Resp. 6. Neither party explains why its proposed definition is appropriate.

Based on the record presented, including our review of the ’196 patent and the types of problems and solutions described in the ’196 patent and cited prior art, we agree with Petitioner’s proposed definition of the level of

ordinary skill in the art and apply it for purposes of this Decision, except that we remove the term “at least” to eliminate vagueness as to the amount of education and practical experience. We note, however, that our conclusions herein would be the same regardless of which party’s definition is used.

C. Claim Interpretation

We interpret each challenged claim

using the same claim construction standard that would be used to construe the claim in a civil action under 35 U.S.C. 282(b), including construing the claim in accordance with the ordinary and customary meaning of such claim as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent.

37 C.F.R. § 42.100(b). Petitioner does not propose express interpretations of any claim terms and states that “[t]he prior art relied on [in the Petition] discloses the subject matter of the challenged claims under any reasonable construction, including their plain meaning.” Pet. 14. Petitioner also notes that Patent Owner argued in the district court case that the “processor” limitations in claims 15 and 19 are not means-plus-function limitations under 35 U.S.C. § 112(f). *Id.* at 15 (citing Ex. 1027, 24–28). Petitioner states that it adopts Patent Owner’s district court construction “[s]olely for the purposes of this IPR.” *Id.* Patent Owner likewise does not propose express interpretations of any claim terms, but disputes Petitioner’s statement regarding “any reasonable construction,” arguing that it “harkens back to the broadest reasonable interpretation standard” previously used by the Board. Prelim. Resp. 5–6. We do not understand Petitioner to have applied that standard, however, given that Petitioner never used the term “broadest.” *See* Pet. 14. We apply the claim interpretation standard set forth

in 37 C.F.R. § 42.100(b) and conclude that no express interpretation of any claim term is necessary to determine whether to institute an *inter partes* review in this proceeding. *See Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (“Because we need only construe terms ‘that are in controversy, and *only to the extent necessary* to resolve the controversy,’ we need not construe [a particular claim limitation] where the construction is not ‘material to the . . . dispute.’” (citation omitted) (emphasis added)).

D. Obviousness Ground Based on Li (Claims 1, 8, 15, and 17)

Petitioner contends that claims 1, 8, 15, and 17 are unpatentable over Li under 35 U.S.C. § 103, citing the testimony of Chris Kyriakakis, Ph.D., as support. Pet. 15–32 (citing Ex. 1003). Patent Owner responds, citing the testimony of Durand R. Begault, Ph.D., as support. Prelim. Resp. 8–34 (citing Ex. 2007). We are not persuaded that Petitioner has established a reasonable likelihood of prevailing on this asserted ground as to any of the challenged claims.

1. Li

Li discloses a 3D sound processing application (3DSPA) that “generates a configurable sound field on a graphical user interface using recorded or pre-recorded sound tracks,” “acquires user selections of configurable parameters associated with sound sources from the configurable sound field,” and “dynamically processes the sound tracks using the user selections to generate a configurable 3D binaural sound . . . by processing the sound tracks with . . . HRTFs based on the user selections.” Ex. 1006, code (57). Li describes a method of measuring

HRTFs using a “simulator apparatus configured to simulate an upper body of a human” (shown in Figure 3A) with microphones in each ear canal that detect impulse sounds from a loudspeaker. *Id.* at col. 9, l. 48–col. 14, l. 4, Figs. 1–6.

Li then describes, with reference to Figure 7, a “method for simultaneously generating configurable three-dimensional (3D) sounds using a microphone array system.” *Id.* at col. 14, ll. 5–7. The 3D sound (e.g., music, speech) may be “3D binaural sound” and may be presented with or without “3D images, 3D movies, and 3D videos.” *Id.* at col. 14, ll. 7–11. The method provides a “3D sound processing application on a computing device” (e.g., laptop) (step 102) and provides a “microphone array system compris[ing] an array of microphone elements . . . configured to form multiple acoustic beam patterns pointing to different positions of multiple sound sources in . . . 3D space” (step 701). *Id.* at col. 14, ll. 19–40. The sound sources may be “similar or different sound generating devices or sound emitting devices, for example, musical instruments, loudspeakers, televisions, music systems, home theater systems, theater systems, a person’s voice, pre-recorded multiple sound tracks, pre-recorded stereo sound tracks, etc.” *Id.* at col. 14, ll. 40–46.

The microphone array system “records sound tracks from the acoustic beam patterns” at the different directions (step 702) and the 3D sound processing application generates a “configurable sound field on a graphical user interface (GUI),” i.e., a “graphical simulation of the source sources in the 3D space” (step 703). *Id.* at col. 14, l. 58–col. 15, l. 3. The GUI “allow[s] the user the freedom to configure the positions and movements of sound sources, in order to generate customized 3D sound.” *Id.* at col. 15, ll. 29–32. The 3D sound processing application allows the user to select,

using the GUI, among “configurable parameters associated with the sound sources,” such as location, azimuth, distance, volume, and “trace of movement” (step 704). *Id.* at col. 15, ll. 32–42. For example, the user can “place the sound sources in specific locations” or “dynamically move the sound sources.” *Id.* at col. 15, ll. 44–48, col. 27, ll. 15–30, Figs. 30A–30B. The 3D sound processing application then dynamically processes “the recorded sound tracks using the acquired user selections to generate one or more of a configurable 3D binaural sound, a configurable 3D surround sound, and a configurable 3D stereo sound.” *Id.* at col. 15, ll. 48–52. For example, the application may process the recorded sound tracks using the HRTFs measured using the simulator apparatus described earlier. *Id.* at col. 15, l. 53–col. 16, l. 11.

Figure 9 of Li is reproduced below.

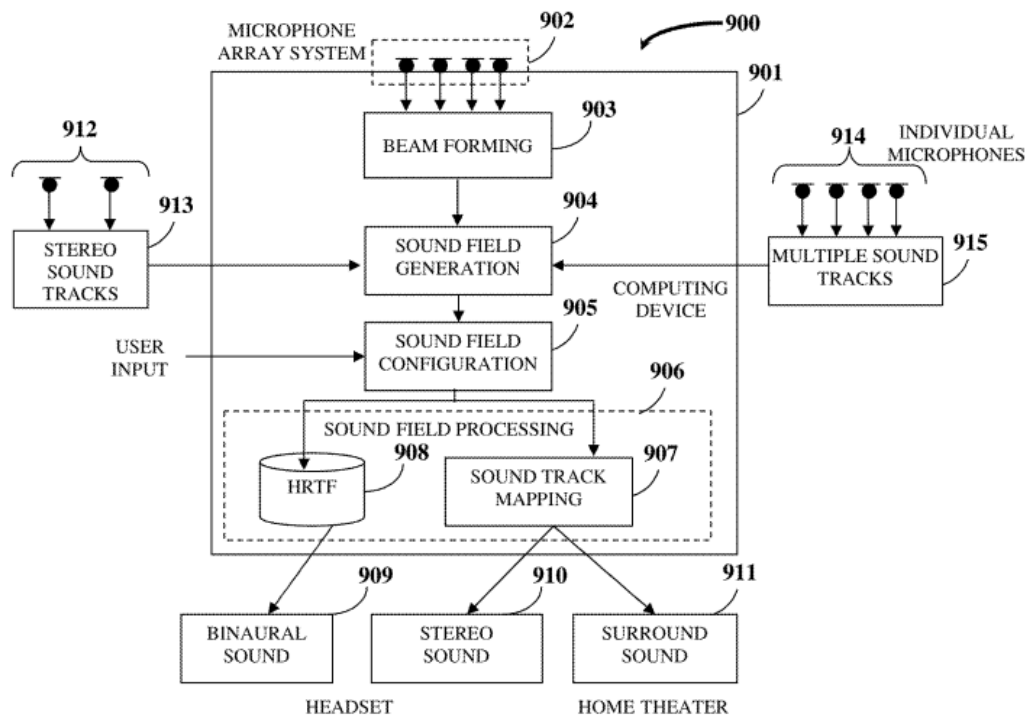


FIG. 9

Figure 9 is a process flow diagram for the steps described above. *Id.* at col. 16, ll. 58–61. Sound tracks are obtained from microphone array system 902, dual microphones 912, or individual microphones 914. *Id.* at col. 17, ll. 41–42, 60–63. “The 3D sound processing application is configured to generate 904, configure 905, and process 906 the configurable sound field.” *Id.* at col. 17, ll. 4–6. The 3D sound processing application then, for example, “performs sound track mapping 907 by convolving each of the sound tracks with corresponding HRTFs stored in the HRTF database 908 to synthesi[ze] 3D binaural sound 909 for a headset user.” *Id.* at col. 17, ll. 33–37.

Figures 17 and 19 of Li are reproduced below.

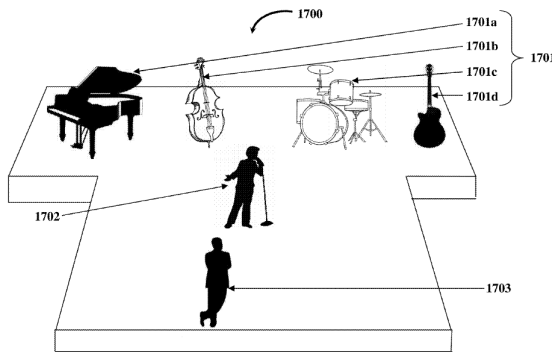


FIG. 17

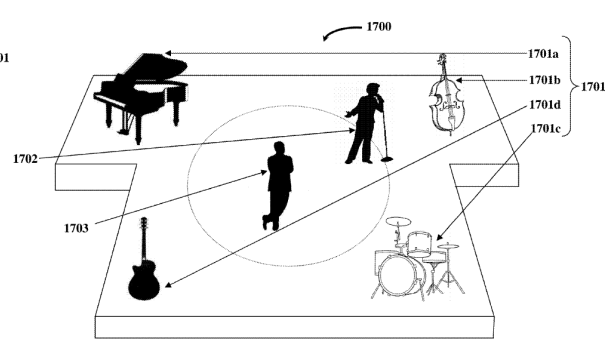


FIG. 19

Figure 17 depicts configurable sound field 1700 that is “a reconstruction of a scene of a concert stage at a music concert” with musical instruments 1701 (piano 1701a, cello 1701b, drums 1701c, and guitar 1701d) and singer 1702; Figure 19 depicts configurable sound field 1700 after user 1703 has altered the positions of the musical instruments and singer using the GUI of the 3D sound processing application. *Id.* at col. 20, l. 53–col. 21, l. 9, col. 21, ll. 38–57. Li discloses that

the configurable 3D sound system 900 disclosed herein allows music artists to present their music to the user in an enhanced manner, and also enhances the performance of radio dramas and

conference calls. The 3D binaural sound recording performed by the configurable 3D sound system 900 disclosed herein provides special effects and acoustic experiences to a user, for example, by allowing the user to move the sound source around the user 1703, move the sound source up and down, etc., on the configurable sound field 1700. The configurable 3D sound system 900 enhances the dramatic performance of radio drama shows and podcasts. Moreover, the configurable 3D sound system 900 provides a method of communication among multiple people, for example, in a conference call, by placing different speaking users at different spots to mimic a real conference room environment.

Id. at col. 21, l. 57–col. 22, l. 5.

2. Claim 1

Petitioner argues that Li teaches all limitations of claim 1 other than the “voice of a computer program,” and asserts that a person of ordinary skill in the art would have been motivated to modify Li to include that feature. Pet. 15–29 (emphasis omitted).

The preamble of claim 1 recites “[a] method executed by one or more electronic devices in an electronic system that move a sound localization point (SLP) of a voice of a computer program during a voice exchange between the computer program and a user.” Petitioner argues that Li’s system “move[s]” a SLP because it allows the user to move sound sources and then generates a configurable sound field based on that input. *Id.* at 21–23. Petitioner contends that “[a]lthough Li does not expressly disclose a voice of a computer program as part of its conference call embodiment, Li teaches using such voice in relation to other embodiments,” pointing to Li’s descriptions of “sound generating devices or sound emitting devices” and sound sources of “a person’s voice” or “pre-recorded multiple sound tracks.”

Id. at 17, 23 (quoting Ex. 1006, col. 4, ll. 58–67, col. 14, ll. 40–48) (emphasis omitted). According to Petitioner, implementing those sound sources as the “voice of a computer program” “would have been advantageous because it would have saved time, cost, and storage space, added flexibility to vocal range, and improved clarity.” *Id.* at 23. Thus, Petitioner appears to advance a first position that the “processed voice” of a human speaker (e.g., a singer) constitutes a “voice of a computer program.” *See id.* at 16–17, 23–24. Petitioner, however, also includes a discussion of “computerized text-to-speech synthesis” that allegedly would have been advantageous as well because, for example, it “would have allowed a speaking-impaired person to participate in Li’s conference call system” and would “add[] flexibility by removing the physical constraints of human speech and adapting to different languages, accents, emotions, and vocal ranges.” *Id.* at 16. Thus, it is unclear from the Petition exactly what modification Petitioner is proposing an ordinarily skilled artisan would have made to Li—either changing the conference call “embodiment” to include processed voice of a speaking user as described elsewhere in Li, or changing it to use text-to-speech computer-synthesized voice rather than the actual voices of the participants, or some combination of both. *See id.* at 15–18, 23–24.

Petitioner further asserts that moving the SLP in Li takes place “during a voice exchange between the computer program and a user” because “[w]hen Li’s user ‘plac[es] different speaking users at different spots to mimic a real conference room environment,’ . . . this takes place during a ‘conference call,’ which involves a voice exchange between the processed voice of the speaker (computer program) and a user/person.” *Id.* at 23–24 (quoting Ex. 1006, col. 21, l. 66–col. 22, l. 5; citing Ex. 1003

¶ 104) (second alteration in original) (emphasis omitted). Dr. Kyriakakis testifies as follows:

I read Li as disclosing that moving the sound source (*mov[ing] a SLP*) takes place *during a voice exchange between the computer program and a user/person*. When Li's user (*user/person*) "plac[es] different speaking users at different spots to mimic a real conference room environment" ([Ex. 1006], 21:66–22:5), this takes place *during* a "conference call," which involves *a voice exchange between the* processed voice of the speaker (*computer program*) *and a user/person*. Li further explains that rather than merely establishing a SLP location before a voice exchange begins, the user (*user/person*) is able to actively move the SLP which would occur after an initial placement has occurred [and] the voice exchange has begun. *Id.*, 21:60–66.

Ex. 1003 ¶ 104 (first two alterations in original).

The body of claim 1 also includes a number of limitations pertaining to the recited "voice exchange," but when addressing those limitations, Petitioner does not provide any additional explanation for how Li allegedly teaches a "voice exchange." With respect to limitation 1[a], which recites "providing . . . binaural sound to the user during the voice exchange between the computer program and the user such that an origin of the voice of the computer program occurs at the SLP at a first location," Petitioner argues that Li's system provides binaural sound to the user over a headset and localizes the SLP "at a location where no physical object exists," and a person of ordinary skill in the art "would have found it obvious to seat the user/person at a physical conference table, where an empty seat at the table . . . is allocated for a virtual participant." Pet. 24–26 (emphasis omitted). With respect to limitation 1[c], which recites "moving, based on the command [to move the SLP received from the user] and during the voice exchange between the computer program and the user, the SLP of the voice

of the computer program from the first location . . . to the second location,” Petitioner points to Li’s disclosure of allowing the user to move sound sources to different locations. *Id.* at 27–28. With respect to limitation 1[d], which recites “providing . . . the binaural sound to the user during the voice exchange between the computer program and the user such that the origin of the voice of the computer program occurs at the SLP at the second location,” Petitioner references its earlier analysis and contends that, after the user has moved a sound source, “the spots for the speakers have changed to new locations” in the sound field. *Id.* at 29. Petitioner does not further address the “voice exchange” aspects of limitation 1[a], 1[c], and 1[d]. *See id.* at 24–29.

Assuming (without deciding) that Li renders obvious a “voice of a computer program” as Petitioner contends,¹ we are not persuaded that Petitioner has made a sufficient showing that Li teaches or suggests a “voice exchange between [a] computer program and [a] user,” or moving a SLP of the voice of a computer program during such a “voice exchange,” as recited in limitations 1[pre],² 1[a], 1[c], and 1[d].

¹ Patent Owner disputes whether Li teaches or suggests a “voice of a computer program,” and also argues that the Petition should be denied under 35 U.S.C. § 314(a) and *Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11 (PTAB Mar. 20, 2020) (precedential). Prelim. Resp. 13–22, 50–59. In light of our decision to deny institution for the reasons discussed herein, however, we need not address these arguments.

² On the record before us, we determine that the preamble is limiting at least because the body of the claim refers to the preamble for antecedent basis for numerous terms, including the “sound localization point (SLP),” “computer program,” “voice of a computer program,” “voice exchange,” and “user.” *See Eaton Corp. v. Rockwell Int’l Corp.*, 323 F.3d 1332, 1339 (Fed. Cir. 2003) (“When limitations in the body of the claim rely upon and derive antecedent basis from the preamble, then the preamble may act as a

Patent Owner argues, and we agree, that “Petitioner fails to explain how . . . Li’s listener can have voice exchange with a computer program” and fails to “identify what features of Li would need to be modified to include [the recited ‘voice exchange’] and explain how such modifications would work.” Prelim. Resp. 22–26. The plain meaning of “voice exchange between [a] computer program and a user,” as recited in claim 1, is an *exchange* of voice audio between the computer program and user. *See id.* at 22; Ex. 1001, col. 35, l. 64–col. 36, l. 1 (the ’196 patent describing an exemplary embodiment where “second user 2362 is *speaking to and listening to* . . . four others” and “the others are not people but computer programs with interactive audio input/output interfaces” (emphasis added)).

Under Petitioner’s first apparent theory that the “voice of a computer program” can be the processed voice of a speaking user as in “other embodiments” of Li, the “other embodiments” all involve *recording* sound tracks and providing a GUI that allows the user to configure the sound that will later be played back. *See, e.g.,* Pet. 17, 23; Prelim. Resp. 21, 24–25; Ex. 1006, code (57) (“The 3DSPA generates a configurable sound field on a graphical user interface using recorded or pre-recorded sound tracks.”), col. 5, ll. 7–26 (“The 3D sound processing application generates a configurable sound field on a graphical user interface (GUI) provided by the 3D sound processing application using the recorded sound tracks,” “acquires user selections of one or more of multiple configurable parameters,” and “dynamically processes the recorded sound tracks using the acquired user selections to generate a configurable 3D binaural sound.”), col. 14,

necessary component of the claimed invention.”). Even if it was not, though, limitations 1[a], 1[c], and 1[d] recite the “voice exchange” discussed herein.

l. 58–col. 15, l. 1 (“The microphone array system records 702 sound tracks from the acoustic beam patterns. . . . The 3D sound processing application generates 703 a configurable sound field . . . using the recorded sound tracks.”), Figs. 7 (steps 702 and 703), 8 (steps 801 and 703), 20 (steps 2001, 2002, and 703), 25 (steps 2501, 2502, and 703), 27 (steps 702 and 703).

Petitioner and Dr. Kyriakakis rely on a single sentence in Li that “the configurable 3D sound system 900 provides a method of communication among multiple people, for example, in a conference call, by placing different speaking users at different spots to mimic a real conference room environment.” Ex. 1006, col. 22, ll. 1–5; *see* Pet. 24; Ex. 1003 ¶ 104.

Petitioner does not provide any further explanation as to how a voice *exchange* would occur during such a “conference call” when modified to use the processed voice of a speaker—which the remainder of Li describes as recorded in advance in sound tracks. Absent additional explanation, we are not persuaded that Li’s brief reference to a “conference call” is sufficient.

Under Petitioner’s second apparent theory that the “voice of a computer program” can be text-to-speech computer-synthesized voice, “mere text-reading with a computer synthesized voice does not constitute a two-way communication” for an *exchange*, and Petitioner includes no explanation of how a modification to Li to use text-to-speech computer-synthesized voice would work. *See* Prelim. Resp. 20, 23–24; Pet. 16.

Moreover, though, Petitioner’s showing is insufficient because the preamble and limitation 1[c] recite “moving” the SLP of the voice of the computer program “*during* the voice exchange between the computer program and the user” (emphasis added). As Patent Owner correctly points out,

Li repeatedly discloses that the sound being moved is “recorded or pre-recorded sound tracks.” None of the examples in Li includes moving a voice during a voice exchange. Instead, all examples and embodiments in Li are directed to recording, and then moving a 3D sound. Thus, in Li no moving of a sound can happen before first recording it

See Prelim. Resp. 26–28 (also citing the description of Figure 19 shown above as a “reconstruction” of a concert involving sound from multiple instruments); Ex. 1006, code (57), col. 4, ll. 7–16 (stating that the configurable 3D binaural sound is generated “using pre-recorded sound tracks and pre-recorded stereo sound tracks”), col. 5, ll. 7–26, col. 14, l. 58–col. 15, l. 3, col. 21, ll. 38–42, Figs. 7, 8, 20, 25, 27; Ex. 2007 ¶ 27 (“Li gives examples using pre-recorded sounds in fixed positions; this method could not be used for a user to interact dynamically in three dimensions with the spatial configuration of one or more sound sources such as computer voices.”). Also,

Li’s users must first input their desired configuration before any binaural sound being rendered—but Li does not disclose a feature that allows users to change their desired configuration during rendering. . . . Li’s music listening example clearly illustrates that users may only place sound sources at their desired spots prior to rendering. . . .

Similarly, Li’s conference call example at best suggests that a user can “plac[e] different speaking users at different spots” only prior to rendering the pre-recorded voices of speaking users at the selected spots—once rendered, Li does not disclose that its users may reconfigure *during their voice exchange* with other speaking users.

See Prelim. Resp. 29–30 (citing Ex. 1006, col. 22, ll. 64–67 (“The configurable 3D sound system 900 renders the 3D binaural sound *from the input configuration* to provide the user with the reconstructed virtual audio 3D space he/she designed.” (emphasis added)), col. 15, l. 61–col. 16, l. 4).

Finally, we agree with Patent Owner that Dr. Kyriakakis’s opinion that he “read[s] Li as disclosing that moving the sound source . . . takes place during a voice exchange between the computer program and a user/person” is conclusory and unsupported. *Id.* at 10–13 (quoting Ex. 1003 ¶ 104) (emphasis omitted). Again, the only support for Dr. Kyriakakis’s opinion is the very brief “conference call” excerpt in Li, which we find insufficient for the reasons explained above. *See* Ex. 1003 ¶ 104 (citing Ex. 1006, col. 21, l. 60–col. 22, l. 5).

Based on the record before us, Petitioner has not shown sufficiently that Li teaches or suggests a “voice exchange between [a] computer program and [a] user,” and moving a SLP of the voice of a computer program during a “voice exchange,” as recited in limitations 1[pre], 1[a], 1[c], and 1[d]. Accordingly, Petitioner has not shown a reasonable likelihood of prevailing on its assertion that claim 1 is unpatentable over Li.

3. *Claims 8, 15, and 17*

Independent claim 8 recites (emphasis added) a “non-transitory computer readable storage medium storing instructions that cause one or more electronic devices in an electronic system to execute a method that *moves a sound localization point (SLP)* of a voice of a computer program *during a voice exchange* that is between the computer program and a person,” the method comprising “provid[ing], through a wearable electronic device on the person, binaural sound to the person *during the voice exchange* that is between the computer program and the person such that an origin of the voice of the computer program occurs at the SLP” at the first and second locations (limitations 8[pre], 8[a], and 8[d]). Independent claim 15 recites (emphasis added) an “electronic system with one or more electronic devices

that *move a sound localization point (SLP)* of a voice of a computer program *during a voice exchange* between the computer program and a person” comprising a “processor that executes . . . instructions to . . . provide binaural sound to the person through the wearable electronic device *during the voice exchange* between the computer program and the person such that an origin of the voice of the computer program occurs at the SLP at the second location” (limitations 15[pre] and 15[c3]). Claim 17 depends from claim 15. Petitioner argues claims 1 and 8 together and, when addressing claim 15, refers to its arguments regarding claim 1, which we address above. Pet. 18–31; *see supra* Section II.D.2. Petitioner has not shown a reasonable likelihood of prevailing on its assertion that claims 8, 15, and 17 are unpatentable over Li.

*E. Additional Obviousness Grounds Based on Li and Other References
(Claims 4–6, 9–11, 14, 18, and 19)*

Petitioner contends, under 35 U.S.C. § 103, that claims 4, 6, 9, 10, 14, and 18 are unpatentable over Li and Ballas; claim 5 is unpatentable over Li, Ballas, and Sandgren; claim 19 is unpatentable over Li and Sandgren; and claim 11 is unpatentable over Li and Woo. Pet. 32–50. Petitioner does not rely on Ballas, Sandgren, or Woo for the limitations of parent independent claims 1, 8, and 15 pertaining to a “voice exchange.” *See id.* Accordingly, for the reasons explained above, Petitioner has not shown a reasonable likelihood of prevailing on its assertion that claims 4–6, 9–11, 14, 18, and 19 are unpatentable. *See supra* Section II.D.

F. Obviousness Ground Based on Bedingfield (Claims 1, 8, 15, and 17)

Petitioner contends that claims 1, 8, 15, and 17 are unpatentable over Bedingfield under 35 U.S.C. § 103, and Patent Owner responds. *See* Pet. 50–65; Prelim. Resp. 34–50. We are not persuaded that Petitioner has established a reasonable likelihood of prevailing on this asserted ground as to any of the challenged claims.

1. Bedingfield

Bedingfield discloses “[a]n improved telephone or network-based communication system” that “simulate[s] an in-person conference setting.” Ex. 1005, col. 3, ll. 25–27. The disclosed system “provide[s] multi-channel sound to a listener by modifying a single channel signal provided by another participant’s system,” and “assign[s] each participant a virtual position within a virtual sound field” where “[e]ach participant is heard by the listener in multi-channel audio based on the participant’s virtual location in the sound field.” *Id.* at col. 3, ll. 27–34.

Figure 1 of Bedingfield is reproduced below.

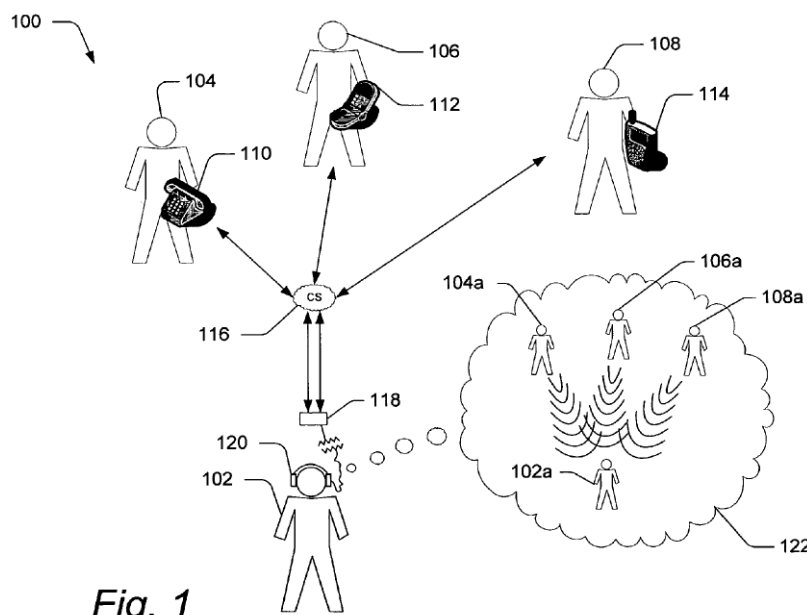


Figure 1 depicts “overall environment 100 for providing a synthetic audio placement to a listener 102” conducting a conference call with participants 104, 106, and 108 having, respectively, communication devices 110, 112, and 114 (e.g., telephones, mobile phones). *Id.* at col. 4, ll. 18–30.

Multi-channel processing device 118 receives signals from the communication devices of speaking participants and creates sound field 122 when listener 102 hears audio through speaker system 120 (e.g., stereo headphones). *Id.* at col. 4, ll. 40–45, col. 13, ll. 34–36. As shown in Figure 1, sound field 122

represents the perception resulting from the audio output of the multi-channel processing device 118 through the multi-channel speaker system 120 as heard by the listener 102. For example, if a listener 102_a and participants 104_a, 106_a, 108_a were in the same location and speaking to each other directly (without the communication system 116), then the listener 102_a would hear sounds from the participants 104_a, 106_a, 108_a through both of the listener’s 102_a ears, and thus have a specific balance, volume, and timbre for each speaker.

Id. at col. 4, ll. 52–61. In that way, created sound field 122 “is substantially similar (or approximately similar) to an actual sound field where the listener 102 is present between all of participants 104, 106, 108.” *Id.* at col. 4, l. 66–col. 5, l. 3, col. 5, ll. 27–32, col. 9, ll. 20–23, col. 16, ll. 36–48, Figs. 19a–c. The sound effects produced for the sound field can be accomplished “using various audio techniques such as volume, balance, parametric equalization, delay, echo and the like.” *Id.* at col. 5, ll. 35–41.

Bedingfield further describes “selections a listener may make” using multi-channel processing device 118. *Id.* at col. 5, ll. 51–53. For example, the listener can “change a participant’s location” by “select[ing] a participant and then select[ing] a new location for the participant within the sound

field.” *Id.* at col. 6, ll. 29–32, 43–45, col. 17, ll. 6–14, 20–25, Fig. 3 (block 328). Bedingfield also discloses detecting movement of the listener (e.g., turning his head) and changing the sound field accordingly for the listener’s new orientation. *Id.* at col. 13, l. 34–col. 14, l. 43, Figs. 13–15.

2. Claim 1

Limitations 1[a] and 1[d] recite (emphasis added):

providing, through an electronic earphone being located at a head of the user and being in communication with a handheld portable electronic device (HPED), *binaural sound* to the user during the voice exchange between the computer program and the user such that an origin of the voice of the computer program occurs at the SLP at a first location in empty space that is away from and proximate to the user; [and]

providing, through the electronic earphone being located at the head of the user and being in communication with the HPED, the *binaural sound* to the user during the voice exchange between the computer program and the user such that the origin of the voice of the computer program occurs at the SLP at the second location in empty space that is proximate to the first location and away from and proximate to the user.

With respect to limitation 1[a], Petitioner argues that Bedingfield teaches providing audio signals to multi-channel speaker system 120 “as headphones” worn by listener 102, shown in Figure 1 above. Pet. 57–58. Petitioner’s only argument regarding the “binaural sound” aspect of the limitation is that “[b]y rendering a sound field and providing synthetic audio placement using both ears, Bedingfield provid[es] binaural sound to the user/person.” *Id.* (emphasis omitted). Dr. Kyriakakis similarly testifies that “[b]y rendering a sound field (which includes multiple SLPs) making use of both ears of a listener (user), Bedingfield discloses providing binaural sound to the user.” Ex. 1003 ¶ 206 (emphasis omitted). With respect to limitation

1[d], Petitioner again argues that Bedingfield teaches providing “the audio rendering of the sound field to both ears (binaural sound) to the listener (user/person),” referencing its earlier analysis of limitation 1[a] and the nearly identical testimony of Dr. Kyriakakis. Pet. 62 (citing *id.* at 57–60; Ex. 1003 ¶ 216) (emphasis omitted). Petitioner and Dr. Kyriakakis give no further explanation as to how Bedingfield teaches providing “binaural sound.”

We are not persuaded that Petitioner has made a sufficient showing for limitations 1[a] and 1[d]. As Patent Owner points out, Petitioner filed a petition in the -1003 case challenging claims of U.S. Patent No. 9,226,090 B1, the parent of the ’196 patent. Prelim. Resp. 44–45. Petitioner argued that “Bedingfield generally discloses all independent claim limitations, except for capturing binaural sound, and maintaining the SLP while the listener moves.” Ex. 2010, 12. Accordingly, Petitioner relied on a secondary reference, U.S. Patent No. 8,818,000 B2 (“Andrea”), arguing that “[w]hile Bedingfield does not expressly teach capturing a binaural sound, this would have been obvious in view of Andrea.” *Id.* at 22. Specifically, Petitioner argued that it would have been obvious to modify Bedingfield’s system to use Andrea’s communication device that has two microphones mounted on the user’s left and right earphones or earbuds (instead of Bedingfield’s communication devices that each have a single microphone). *Id.* at 16–19, 21–22. According to Petitioner, “[m]odifying Bedingfield [based on Andrea] would have been as simple as adding a second microphone to Bedingfield’s WED for capturing binaural voice.” *Id.* at 18. Petitioner does not rely on Andrea, or any other secondary reference, as

teaching “binaural sound” for purposes of its asserted ground based on Bedingfield in this proceeding.³

Bedingfield does not use the term “binaural sound” and Petitioner does not point to any other language in the reference that could be equated with, or suggest, providing binaural sound. *See* Prelim. Resp. 45. Indeed, it is unclear how, if Bedingfield does not *capture* “binaural sound” from participants 104, 106, and 108 (as Petitioner expressly acknowledges in the -1003 case), Bedingfield would *provide* “binaural sound” to listener 102 (as Petitioner argues in this proceeding). Petitioner and Dr. Kyriakakis provide no explanation why that would be the case, other than to state that listener 102’s headphones are at both ears. *See* Pet. 58, 62; Ex. 1003 ¶¶ 206, 216; *see also* IPR2023-01003, Ex. 1003 ¶¶ 139–140 (Dr. Kyriakakis testifying in the -1003 case that “Bedingfield in view of Andrea” teaches “providing . . . binaural sound” because “the audio signals (sound) are provid[ed] to the listening person during the telephone call, and Bedingfield in view of Andrea discloses the sound is binaural sound” (emphasis omitted)).

Further, Dr. Kyriakakis testifies:

Binaural sound can be created in two ways. The first is to record binaural sound using two microphones that are spaced at the same distance as the width of a human head (i.e., a binaural recording). This is often achieved by placing microphones on the sides of a mannequin (or dummy) head. When the left microphone signal is sent to the left ear and the right microphone signal is sent to the right ear (using headphones), the listener

³ In a concurrently entered decision, we determine that Petitioner has shown a reasonable likelihood of prevailing as to certain challenged claims based on the Bedingfield-Andrea combination and institute an *inter partes* review in the -1003 case.

perceives the recorded scene with precise localization of each source in 3-D space.

The second is to synthesize binaural sound by processing non-binaural content with the left and right HRTFs of the desired placement angle and position. For example, a monaural sound source, such as a person's voice, is processed (via convolution) with a left ear HRTF and a right ear HRTF, where each HRTF is designated at the desired placement angle and position, such as 30° to the front left of the listener and at a specified distance from the listener. This processing produces a "virtual auditory space," with a two-channel synthetic binaural signal with the left channel sent to the listener's left ear and the right channel to the listener's right ear (using headphones). At the relevant time period, HRTFs had been known for decades and were commonly used to characterize the manner in which an ear receives a sound from a point in space.

Ex. 1003 ¶¶ 52–53 (citations and emphasis omitted). We agree with Patent Owner that Bedingfield does not teach either method. *See* Prelim. Resp. 45–47. In any case, Petitioner does not explain in any detail *how* Bedingfield allegedly teaches or suggests "binaural sound" provided to the listener who is communicating with other participants on a conference call.⁴

Accordingly, Petitioner has not shown a reasonable likelihood of prevailing on its assertion that claim 1 is unpatentable over Bedingfield.

3. *Claims 8, 15, and 17*

Independent claim 8 recites (emphasis added) a "non-transitory computer readable storage medium storing instructions that cause one or

⁴ Patent Owner also argues that Bedingfield does not teach or suggest a "voice of a computer program" or "voice exchange between [a] computer program and a user," as recited in claim 1 and the other independent claims. Prelim. Resp. 35–43, 47–50. In light of our decision to deny institution for the reasons discussed herein, however, we need not address these arguments.

more electronic devices in an electronic system to execute a method” comprising “provid[ing], through a wearable electronic device on the person, *binaural sound* to the person during the voice exchange that is between the computer program and the person” (limitations 8[a] and 8[d]). Independent claim 15 recites (emphasis added) an “electronic system” comprising a “processor that executes . . . instructions to . . . provide *binaural sound* to the person through the wearable electronic device during the voice exchange between the computer program and the person” (limitation 15[c3]). Claim 17 depends from claim 15. Petitioner argues claims 1 and 8 together and, when addressing claim 15, refers to its arguments regarding claim 1, which we address above. Pet. 57–60, 62, 64–65; *see supra* Section II.F.2. Petitioner has not shown a reasonable likelihood of prevailing on its assertion that claims 8, 15, and 17 are unpatentable over Bedingfield.

G. Additional Obviousness Grounds Based on Bedingfield and Other References (Claims 4–6, 9–11, 14, 18, and 19)

Petitioner contends, under 35 U.S.C. § 103, that claims 4, 6, 9, 10, 14, and 18 are unpatentable over Bedingfield and Ballas; claim 5 is unpatentable over Bedingfield, Ballas, and Sandgren; claim 19 is unpatentable over Bedingfield and Sandgren; and claim 11 is unpatentable over Bedingfield and Woo. Pet. 65–79. Petitioner does not rely on Ballas, Sandgren, or Woo for the limitations of parent independent claims 1, 8, and 15 pertaining to “binaural sound.” *See id.* Accordingly, for the reasons explained above, Petitioner has not shown a reasonable likelihood of prevailing on its assertion that claims 4–6, 9–11, 14, 18, and 19 are unpatentable. *See supra* Section II.F.

III. CONCLUSION

Based on the arguments and evidence presented by the parties, we conclude that Petitioner has not demonstrated a reasonable likelihood of prevailing with respect to at least one claim of the '196 patent challenged in the Petition. Therefore, we do not institute an *inter partes* review.

IV. ORDER

In consideration of the foregoing, it is hereby:

ORDERED that the Petition is denied and no *inter partes* review is instituted.

FOR PETITIONER:

Lisa K. Nguyen
Grace I. Wang
David A. Hubbard
Alan M. Billharz
David M. Tennant
ALLEN & OVERY LLP
lisa.nguyen@allenovery.com
grace.wang@allenovery.com
david.hubbard@allenovery.com
alan.billharz@allenovery.com
david.tennant@allenovery.com

FOR PATENT OWNER:

Kevin S. Tuttle
Erick S. Robinson
Patrick M. Dunn
Kyle L. Elliott
Homayoon Rafatijo
SPENCER FANE LLP
ktuttle@spencerfane.com
erobinson@spencerfane.com
patdunn64@gmail.com
kelliott@spencerfane.com
hrafatijo@spencerfane.com