

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

LG ELECTRONICS, INC.,
Petitioner,

v.

JAWBONE INNOVATIONS, LLC,
Patent Owner.

IPR2023-01110
Patent 7,246,058 B2

Before GEORGIANNA W. BRADEN, NORMAN BEAMER, and
JASON M. REPKO, *Administrative Patent Judges*.

BRADEN, *Administrative Patent Judge*.

DECISION
Granting Institution of *Inter Partes* Review
35 U.S.C. § 314
Denying Motion for Joinder
35 U.S.C. § 315(c); 37 C.F.R. § 42.122

I. INTRODUCTION

LG Electronics, Inc. (“Petitioner”) filed a Petition requesting an *inter partes* review of claims 1–5 in U.S. Patent No. 7,246,058 B2 (Exhibit 1001, “the ’058 patent”) under 35 U.S.C. §§ 311–319. Paper 2 (“Pet.”). Jawbone Innovations, LLC (“Patent Owner”) filed a Preliminary Response. Paper 9 (“Prelim. Resp.”).

Under 37 C.F.R. § 42.4(a), we have authority to determine whether to institute an *inter partes* review. We may institute an *inter partes* review only if “the information presented in the petition filed under section 311 and any response filed under section 313 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.” 35 U.S.C. § 314(a) (2018). The “reasonable likelihood” standard is “a higher standard than mere notice pleading” but “lower than the ‘preponderance’ standard to prevail in a final written decision.” *Hulu, LLC v. Sound View Innovations, LLC*, IPR2018-01039, Paper 29 at 13 (PTAB Dec. 20, 2019) (precedential).

Based on the current record and for the reasons explained below, Petitioner has shown that there is a reasonable likelihood that it would prevail with respect to at least one of the challenged claims. Thus, we institute an *inter partes* review of claims 1–5 in the ’058 patent on all challenges included in the Petition.

II. BACKGROUND

A. *Real Parties in Interest*

Petitioner identifies LG Electronics, Inc. as the real party in interest. Pet. 62. Patent Owner identifies itself as the real party in interest. Paper 6, 2. The parties do not raise any issue about real parties in interest.

B. Related Matters

The parties both identify the following multiple district court cases involving the '058 patent as well as related IPR proceedings: IPR2022-00623; IPR2022-01021; IPR2023-00222; and IPR2023-01119. Pet. 63–64; Paper 6, 2–3. Each of the related IPR proceedings, except IPR2023-01119, has been terminated due to settlement. *See* IPR2022-00623, Paper 31; IPR2022-01021, Paper 14; IPR2023-00222, Paper 13.

C. Motion for Joinder and Consolidation

Petitioner moves for joinder to and consolidation with related IPR proceeding IPR2023-00222 under 35 U.S.C. § 315 and 37 C.F.R. § 42.122(b). Paper 3 (Motion for Joinder). IPR proceeding IPR2023-00222 has been terminated due to settlement. *See* IPR2023-00222, Paper 13. Thus, there is no proceeding with which to join the present case. Accordingly, Petitioner's motion for joinder and consolidation is denied.

D. The '058 Patent (Exhibit 1001)

The '058 patent, titled “Detecting Voiced and Unvoiced Speech Using Both Acoustic and Nonacoustic Sensors,” issued on July 17, 2007, from an application filed on May 30, 2002. Ex. 1001, codes (22), (45), (54). The patent claims priority to the following applications: (1) eleven provisional applications filed between May 30, 2001, and March 27, 2002; (2) a nonprovisional application filed on July 12, 2001; and (3) a nonprovisional application filed on November 21, 2001. *Id.* at 1:7–14, code (60).

1. Written Description

The '058 patent discloses systems and methods for “detecting voiced and unvoiced speech in acoustic signals having varying levels of background noise.” *Id.* at code (57).

The '058 patent's Figures 1 and 2 (reproduced below) depict embodiments of block diagrams of speech-detection systems 100 and 200:

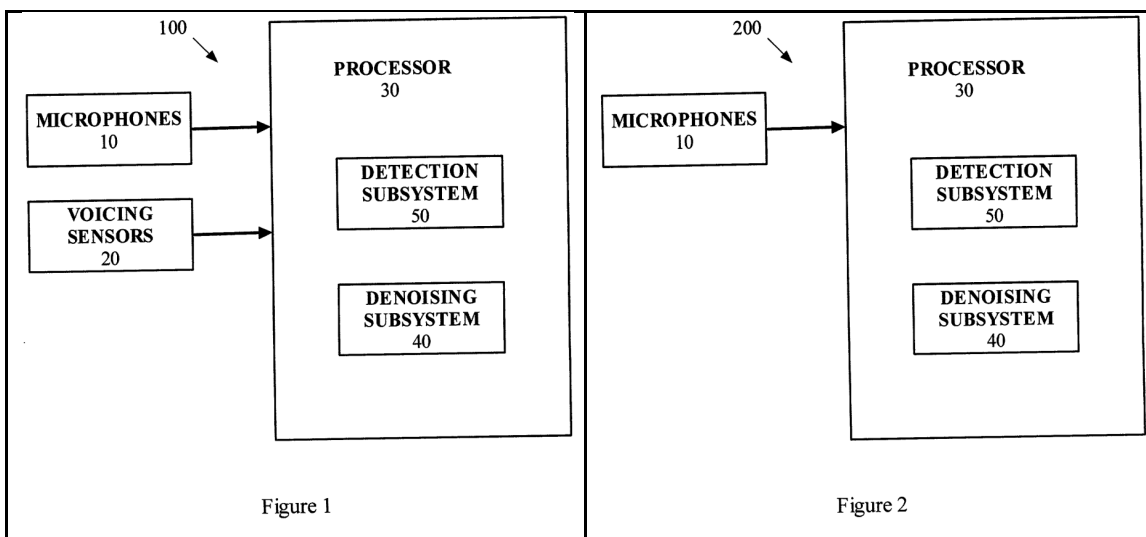


Figure 1 (above on the left) illustrates “a Non-Acoustic Sensor Voiced Speech Activity Detection (NAVSAD) system” 100, and Figure 2 (above on the right) illustrates “a Pathfinder Speech Activity Detection (PSAD) system” 200. Ex. 1001, 1:52–55, 2:20–24, 2:30–31, 2:41–42, Figs. 1–2. “The PSAD can operate independently or as a backup to the NAVSAD, detecting voiced speech if the NAVSAD fails.” *Id.* at 2:48–49; *see id.* at 8:62–63, 9:44–45.

As Figures 1 and 2 show, both NAVSAD system 100 and PSAD system 200 include microphones 10 coupled to processor 30. Ex. 1001, 2:31–32, 2:42–43, Figs. 1–2. NAVSAD system 100 also includes voicing sensors 20, i.e., nonacoustic sensors, coupled to processor 30. *Id.* at 2:31–34, Fig. 1. Processor 30 controls subsystems including denoising

subsystem 40 and detection subsystem 50, e.g., a detection algorithm. *Id.* at 2:34–37, 2:43–45; *see id.* at 2:50–60.

The '058 patent explains that unvoiced speech detection should be:

- (1) “insensitive” in high-noise situations because unvoiced speech is “easily washed out by the noise” and “unlikely to affect the voiced speech denoising”; and
- (2) “highly sensitive” in low-noise situations because the “speech information is most important in the presence of little to no noise.”

Ex. 1001, 3:17–27; *see id.* at 5:65–6:1.

The '058 patent also explains that NAVSAD system 100 includes a nonacoustic sensor “to enable detection of voiced speech in any environment” by detecting “human tissue motion associated with speech.” Ex. 1001, 3:13–15, 3:55–63. As an example of a nonacoustic sensor, the patent identifies a General Electromagnetic Movement Sensor (GEMS), i.e., “a radio frequency device (2.4 GHz) that allows the detection of moving human tissue dielectric interfaces.” *Id.* at 3:57–60, 3:64–66. “The GEMS includes an RF interferometer that uses homodyne mixing to detect small phase shifts associated with target motion.” *Id.* at 3:63–4:1.

“The NAVSAD and PSAD systems include an array algorithm for speech detection that uses the difference in frequency content between two microphones to calculate a relationship between the signals of the two microphones.” Ex. 1001, 3:29–32.

The '058 patent's Figure 4 (reproduced below) depicts a flow diagram of a detection algorithm for detecting voiced and unvoiced speech:

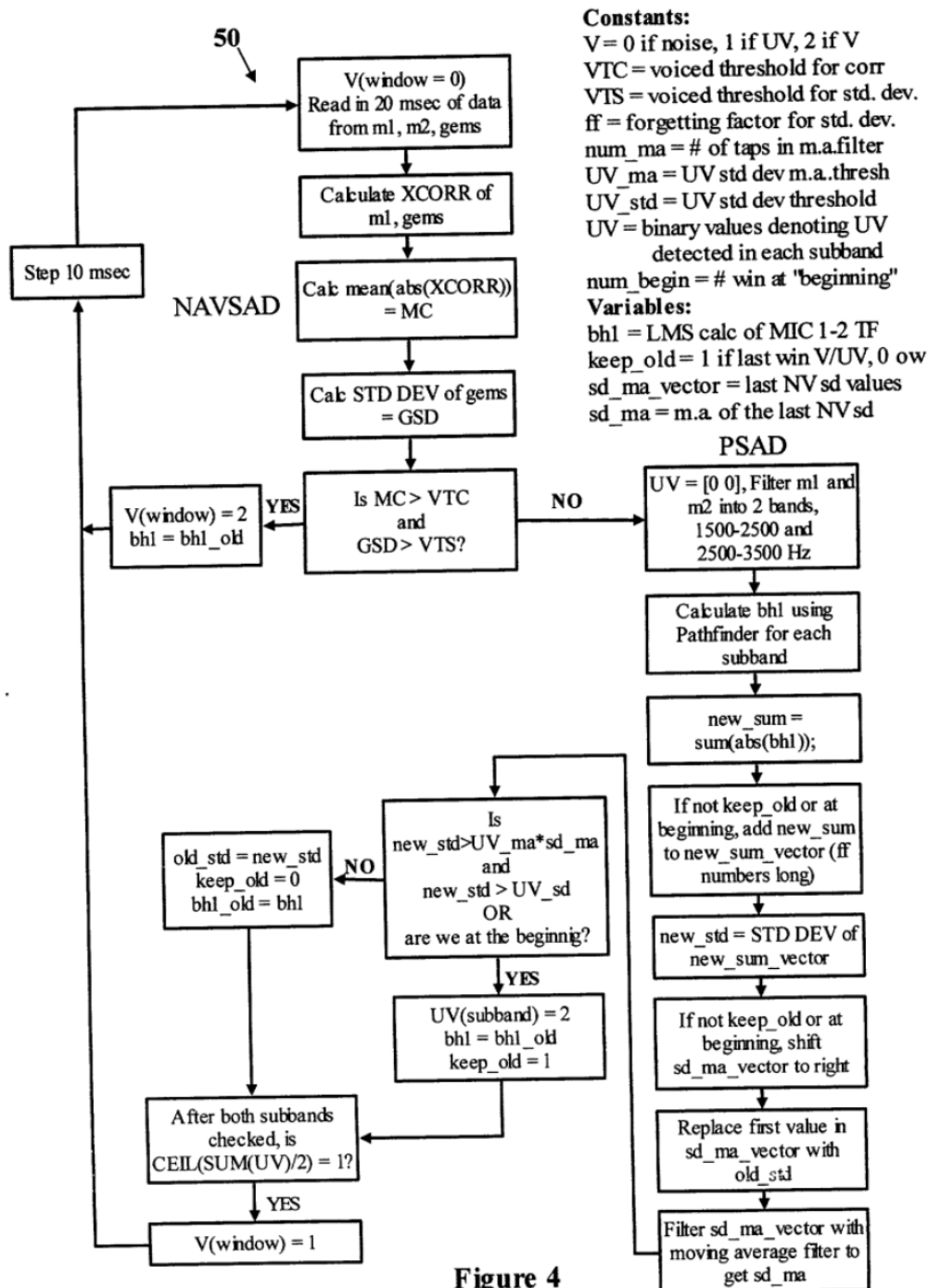


Figure 4 illustrates detection algorithm 50 for detecting voiced and unvoiced speech in “environments containing varying amounts of background

acoustic noise.” Ex. 1001, 1:58–60, 4:14–16, 4:42–44, Fig. 4. NAVSAD system 100 and PSAD system 200 may “include the detection algorithm 50 as the detection subsystem 50.” *Id.* at 4:16–18.

A system with detection algorithm 50 records data from a first microphone (“m1” in Figure 4), a second microphone (“m2” in Figure 4), and a nonacoustic sensor (“gems” in Figure 4) during a time window, e.g., 20 milliseconds. Ex. 1001, 4:18–26, Fig. 4; *see id.* at code (57). The system determines the energy of the signals from the first microphone, the second microphone, and the nonacoustic sensor over the time window. *Id.* at 4:64–5:3.

The system generates “difference parameters between the acoustic signals received at each of the two microphones.” Ex. 1001, code (57); *see id.* at 7:52–8:48. “The difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals.” *Id.* at code (57). The system compares the difference parameters to various thresholds to distinguish speech from noise. *Id.*

The system also determines the cross-correlation (XCORR) between the acoustic signal from the first microphone and the signal from the nonacoustic sensor to “further differentiate sensor noise from tissue motion.” Ex. 1001, 4:64–5:3, 5:11–12, Fig. 4; *see id.* at 5:26–44. “The sensor detects human tissue motion associated with the closure of the vocal folds, so the acoustic signal produced by the closure of the folds is highly correlated with the closures.” *Id.* at 5:26–29. Hence, “sensor data that correlates highly with the acoustic signal is declared as speech, and sensor data that does not correlate well is termed noise.” *Id.* at 5:29–31.

2. Illustrative Claim

Petitioner challenges independent system claim 1, independent method claim 2, and claims 3–5 that depend directly from claim 2. Pet. 1, 18–56.

Claim 1 exemplifies the challenged claims and reads as follows (with formatting added for clarity):

1. A system for detecting voiced and unvoiced speech in acoustic signals having varying levels of background noise, comprising:

at least two microphones that receive the acoustic signals;

at least one voicing sensor that receives physiological information associated with human voicing activity; and

at least one processor coupled among the microphones and the voicing sensor,

wherein the at least one processor;

generates cross correlation data between the physiological information and an acoustic signal received at one of the two microphones;

identifies information of the acoustic signals as voiced speech when the cross correlation data corresponding to a portion of the acoustic signal received at the one receiver exceeds a correlation threshold;

generates difference parameters between the acoustic signals received at each of the two receivers, wherein the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals;

identifies information of the acoustic signals as unvoiced speech when the difference parameters exceed a gain threshold; and

identifies information of the acoustic signals as noise when the difference parameters are less than the gain threshold.

Ex. 1001, 11:8–35.

E. Evidence of Record

For its challenges, Petitioner relies on the following references:

Name	Reference	Exhibit
Holzrichter	US 6,006,175, issued December 21, 1999 (based on an application filed February 6, 1996)	1003
Kinsler	Lawrence E. Kinsler et al., <i>Fundamentals of Acoustics</i> (3d ed. 1982)	1004
Takano ¹	JP H11-305792, published November 5, 1999 (based on an application filed August 19, 1998)	1005
Torick	E.L. Torick & R.G. Allen, <i>An Interphone System for “Hands-Free” Operation in High Ambient Noise</i> , IEEE Transactions on Audio & Electroacoustics, Vol. AU-14, No. 4, 168–73 (December 1966)	1006
Kaneda	US 5,208,864, issued May 4, 1993 (based on an application filed March 8, 1990)	1007
Lyons	Richard G. Lyons, <i>Understanding Digital Signal Processing</i> (1997)	1008

Pet. 14, 17–56.

Petitioner asserts that each reference qualifies as prior art under § 102(b). Pet. 14–15; *see* 35 U.S.C. § 102(b) (2006).²

¹ Petitioner provides JP H11-305792 (Takano) in Japanese as Exhibit 1018 and an English translation as Exhibit 1005. This decision cites the English translation (Exhibit 1005).

² The Leahy-Smith America Invents Act (“AIA”), Pub. L. No. 112-29, 125 Stat. 284 (2011), amended 35 U.S.C. § 102 and § 103 effective March 16, 2013. Because the filing date of the challenged claims predates

At this stage of the proceeding, Patent Owner does not dispute that each reference qualifies as prior art. *See, e.g.*, Prelim. Resp. 4–17.

F. The Asserted Challenges to Patentability

Petitioner asserts the following challenges to patentability:

Claim(s) Challenged	35 U.S.C. §	Reference(s)/Basis
1–5	103(a)	Holzrichter, Takano, Kinsler
3	103(a)	Holzrichter, Takano, Kinsler, Lyons
1, 2, 4, 5	103(a)	Holzrichter, Torick, Kaneda, Kinsler
3	103(a)	Holzrichter, Torick, Kaneda, Kinsler, Lyons

Pet. 15, 18–56.

G. Testimonial Evidence

To support its challenges, Petitioner relies on the declaration of Richard M. Stern, Ph.D. (Exhibit 1002). Dr. Stern states, “I received a S.B. degree from the Massachusetts Institute of Technology (MIT) in 1970, a M.S. degree from the University of California, Berkeley, in 1972, and a Ph.D. from MIT in 1977, all in electrical engineering.” Ex. 1002 ¶ 5.

Dr. Stern also states, “I am a Professor at Carnegie Mellon University in the Department of Electrical and Computer Engineering, the Department of Computer Science, and the Language Technologies Institute.” *Id.* ¶ 4.

the AIA’s amendments to § 102 and § 103, this decision refers to the pre-AIA versions of § 102 and § 103.

Additionally, he explains that “I have been retained by Amazon as a technical expert in this matter.”³ *Id.* ¶ 1.

Petitioner also relies on the declaration of Carol S. Peterson (Exhibit 1019). Ms. Peterson states, “I earned a Master of Library Science from the University of California Los Angeles in 1979 and a Bachelors in English from the University of California, Davis in 1977. I have worked as a research librarian at Knobbe, Martens, Olson & Bear for approximately 40 years.” Ex. 1019 ¶ 3. Ms. Peterson discusses “standard library practices for making materials available to the public” as relevant to Kinsler, Lyons, and Torick. *Id.* ¶¶ 9–56.

Petitioner additionally relies on the declaration of Brian Toth (Exhibit 1020). Mr. Toth states, “I submit this declaration to certify the accuracy of the English translation of Japanese Patent Application Publication No. H-11-305792,” i.e., Takano. Ex. 1020 ¶ 2.

³ Dr. Stern was retained as an independent expert by non-party Amazon.com, Inc. for IPR proceeding IPR2023-00222, which challenged to the ’058 patent based on the same prior art as asserted in the present Petition. We remind the parties of their obligation under Rule 42.51(b)(1)(ii), which allows for cross examination of affidavit testimony as routine discovery. *See John’s Lone Star Distrib., Inc. v. Thermolife Int’l, Inc.*, Case No. IPR2014–01201 (PTAB May 13, 2015) (Paper 31) (petitioner submitted affidavit testimony from a witness triggering cross-examination under 37 C.F.R. § 42.51(b)(1), and because petitioner did not make that witness available for cross-examination, the Board expunged the testimony.); *IBG LLC et al. v. Trading Technologies Internationale, Inc.*, Case No. CBM2015-00179 (PTAB Apr. 15, 2016) Paper 39, 3 (The Board determined that it was the petitioner’s responsibility to produce the witness for cross-examination, because that deposition was the sole evidence that a reference was prior art.).

III. PATENTABILITY ANALYSIS

A. Legal Principles: Obviousness

A patent may not be obtained “if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains.” 35 U.S.C. § 103(a) (2006). An obviousness analysis involves underlying factual inquiries including (1) the scope and content of the prior art; (2) differences between the claimed invention and the prior art; (3) the level of ordinary skill in the art; and (4) where in evidence, objective indicia of nonobviousness, such as commercial success, long-felt but unsolved needs, and failure of others.⁴ *Graham v. John Deere Co.*, 383 U.S. 1, 17–18, 35–36 (1966); *Apple Inc. v. Samsung Elecs. Co.*, 839 F.3d 1034, 1047–48 (Fed. Cir. 2016) (en banc). When evaluating a combination of references, an obviousness analysis should address “whether there was an apparent reason to combine the known elements in the fashion claimed by the patent at issue.” *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007).

We analyze the obviousness issues according to these principles.

B. Level of Ordinary Skill in the Art

Factors pertinent to determining the level of ordinary skill in the art include (1) the educational level of the inventor; (2) the type of problems encountered in the art; (3) prior-art solutions to those problems; (4) the rapidity with which innovations are made; (5) the sophistication of the technology; and (6) the educational level of workers active in the field.

⁴ The record does not include evidence concerning objective indicia of nonobviousness.

Envtl. Designs, Ltd. v. Union Oil Co., 713 F.2d 693, 696–97 (Fed. Cir. 1983). Not all factors may exist in every case, and one or more of these or other factors may predominate in a particular case. *Id.* These factors are not exhaustive, but merely a guide to determining the level of ordinary skill in the art. *Daiichi Sankyo Co. v. Apotex, Inc.*, 501 F.3d 1254, 1256 (Fed. Cir. 2007). Moreover, the prior art itself may reflect an appropriate skill level. *Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001).

Petitioner asserts that a person of ordinary skill in the art at the time of the alleged invention “would have had a minimum of a bachelor’s degree in computer engineering, computer science, electrical engineering, mechanical engineering, or a similar field, and approximately three years of industry or academic experience in a field related to acoustics, speech recognition, speech detection, or signal processing.” Pet. 16–17. Petitioner also asserts that “[w]ork experience can substitute for formal education and additional formal education can substitute for work experience.” *Id.* at 16. Dr. Stern’s testimony supports Petitioner’s assertions. *See* Ex. 1002 ¶¶ 30–32.

For purposes of the Preliminary Response, Patent Owner “utilizes Petitioner’s proposed level of skill in the art.” Prelim. Resp. 7 (quoting Pet. 165 [*read*: 16]–17).

Based on the current record and for purposes of analysis, we accept Petitioner’s description of an ordinarily skilled artisan as it is consistent with the level of skill shown in the ’058 patent and the asserted prior art.

C. Claim Construction

We construe claim terms “using the same claim construction standard” that district courts use to construe claim terms in civil actions

under 35 U.S.C. § 282(b). *See* 37 C.F.R. § 42.100(b) (2022). Under that standard, claim terms “are given their ordinary and customary meaning, which is the meaning the term would have to a person of ordinary skill in the art at the time of the invention.” *Power Integrations, Inc. v. Fairchild Semiconductor Int’l, Inc.*, 904 F.3d 965, 971 (Fed. Cir. 2018) (citing *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–13 (Fed. Cir. 2005) (en banc)). The meaning of claim terms may be determined by “look[ing] principally to the intrinsic evidence of record, examining the claim language itself, the written description, and the prosecution history, if in evidence.” *DePuy Spine, Inc. v. Medtronic Sofamor Danek, Inc.*, 469 F.3d 1005, 1014 (Fed. Cir. 2006) (citing *Phillips*, 415 F.3d at 1312–17).

Petitioner contends that “no claim terms require construction to resolve the obviousness challenges.” Pet. 17.

Patent Owner “believes that claim construction is not required to resolve any issues.” Prelim. Resp. 6.

Based on the current record, we determine that no claim term requires an explicit construction to decide whether Petitioner satisfies the “reasonable likelihood” standard for instituting trial. “[O]nly those terms need be construed that are in controversy, and only to the extent necessary to resolve the controversy.” *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999); *see Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017). The analysis below, however, discusses the meaning of “noise” as recited in claim 1. *See supra* §§ E.1.i. & F.1.e. In particular, at this stage and on this record, we disagree with Patent Owner’s arguments in claim 1 because we determine that the claim 1 sets the boundary for what is considered “noise” for the reasons

discussed below. *See id.* During trial, the parties are invited to address whether the Board should read the claim in this way in its Final Decision.

D. Overview of Asserted Prior Art of Record

1. Overview of Holzrichter (Exhibit 1003)

Holzrichter is a U.S. patent titled “Methods and Apparatus for Nonacoustic Speech Characterization and Recognition,” filed on February 6, 1996, and issued on December 21, 1999. Ex. 1003, codes (11), (22), (45), (54). Holzrichter states that the “invention relates generally to speech recognition and more particularly to the use of nonacoustic information in combination with acoustic information for speech recognition and related speech technologies.” *Id.* at 1:12–15; *see id.* at 7:21–24, code (57).

More specifically, Holzrichter explains that the “invention is [a] method and apparatus for nonacoustic speech recognition (NASR) in which nonacoustic information obtained by [radio-frequency] RF wave, microwave, millimeter wave, infrared, or optical wave electromagnetic (EM) measurements of speech organs is combined with conventional acoustic information measured with a microphone.” Ex. 1003, 7:37–42. According to Holzrichter, “no one has previously suggested a speech recognition system using transmitted and reflected EM waves to detect motions and locations of speech organs and to use the information in an algorithm to identify speech.” *Id.* at 7:14–18.

Holzrichter's Figure 4 (reproduced below) depicts an apparatus for nonacoustic speech recognition (NASR):

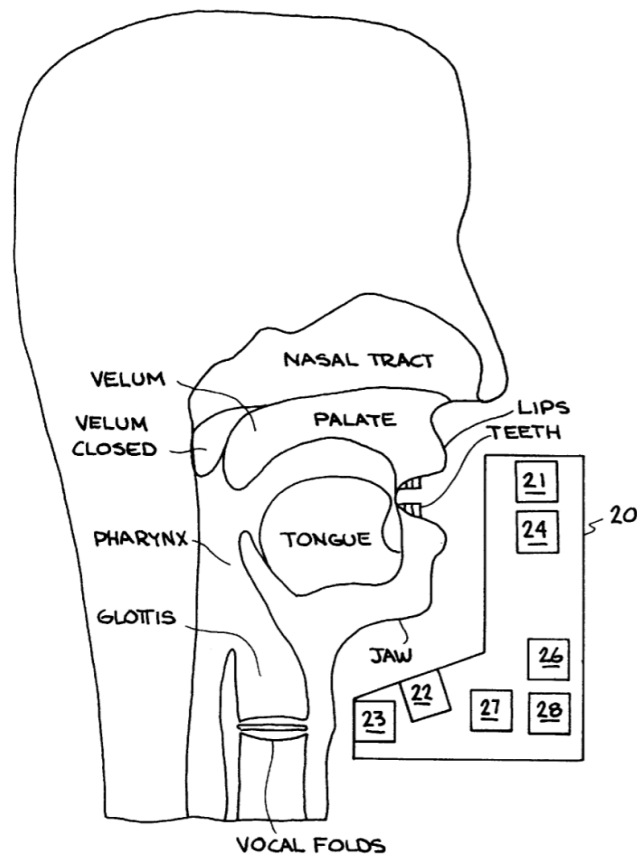


FIG. 4

Figure 4 illustrates NASR sensor 20 and “a cross sectional view of a head showing speech organs” with NASR sensor 20 positioned near the head.

Ex. 1003, 15:11–13, 17:53–54, Fig. 4.

NASR sensor 20 includes EM sensor transmit/receive modules 21, 22, and 23. Ex. 1003, 17:54–56, Fig. 4. Module 21 directs EM waves toward “the lips, teeth and tongue tip.” *Id.* at 17:61–62, Fig. 4. Module 22 directs EM waves toward “the lower jaw, and into the lip palate organ region.” *Id.* at 17:64–65, Fig. 4. Module 23 directs EM waves toward “the vocal folds and glottis.” *Id.* at 17:62–63, Fig. 4. Although Figure 4 shows three EM

sensor transmit/receive modules, “there can be as few as one or more than three.” *Id.* at 17:54–56.

NASR sensor 20 includes microphone 24 “used to obtain simultaneous acoustic information.” Ex. 1003, 17:54–58, Fig. 4. NASR sensor 20 also includes “the necessary electronics, e.g. timing and range gate chip 26 and memory chip 27, as well as means to transfer the data over a wire or wireless, e.g. driver chip 28.” *Id.* at 17:58–61, Fig. 4.

Holzrichter explains that the “techniques of nonacoustic, EM sensor derived speech organ position and motion measurement are statistically independent of acoustical speech measurement using a microphone.” Ex. 1003, 10:26–29. Thus, “the two sets of information complement each other in a statistical measurement sense.” *Id.* at 10:29–30.

Holzrichter discloses correlating EM-sensor data and audio (acoustic). Ex. 1003, 15:49–51, 16:27–29, 20:11–15, 26:44–27:4, 45:2–15, Figs. 15A–15D, 25A–25H. Holzrichter describes one process for correlating EM-sensor data and audio (acoustic) data as follows:

The acoustic signals from a microphone sensing the same words as the EM sensor(s) were displayed on a separate trace of the oscilloscope and were sampled in an A/D converter and were stored in a memory with identical start times and with identical time bin numbers as the memory bins for the EM sensor data. The correlation between these two (or more) signals were used to validate the assignment of irreducible spoken sound units (i.e. [phone-like units] PLUs) with the expected location or activity of the vocal organ being measured.

Ex. 1003, 26:60–27:2.

Holzrichter explains that for “each set of received EM wave signals there is a need to process and obtain the information on organ motions (or their new positions) that can be used to associate the unknown speech with

the intended speech sound or word sound.” Ex. 1003, 9:23–27.

Specifically, the “nonacoustic and acoustic signals are combined using an algorithm to produce more accurate speech recognition than obtainable only from acoustic information.” *Id.* at 7:44–46.

Holzrichter’s Figure 19 (reproduced below) depicts an algorithm for analyzing nonacoustic and acoustic signals:

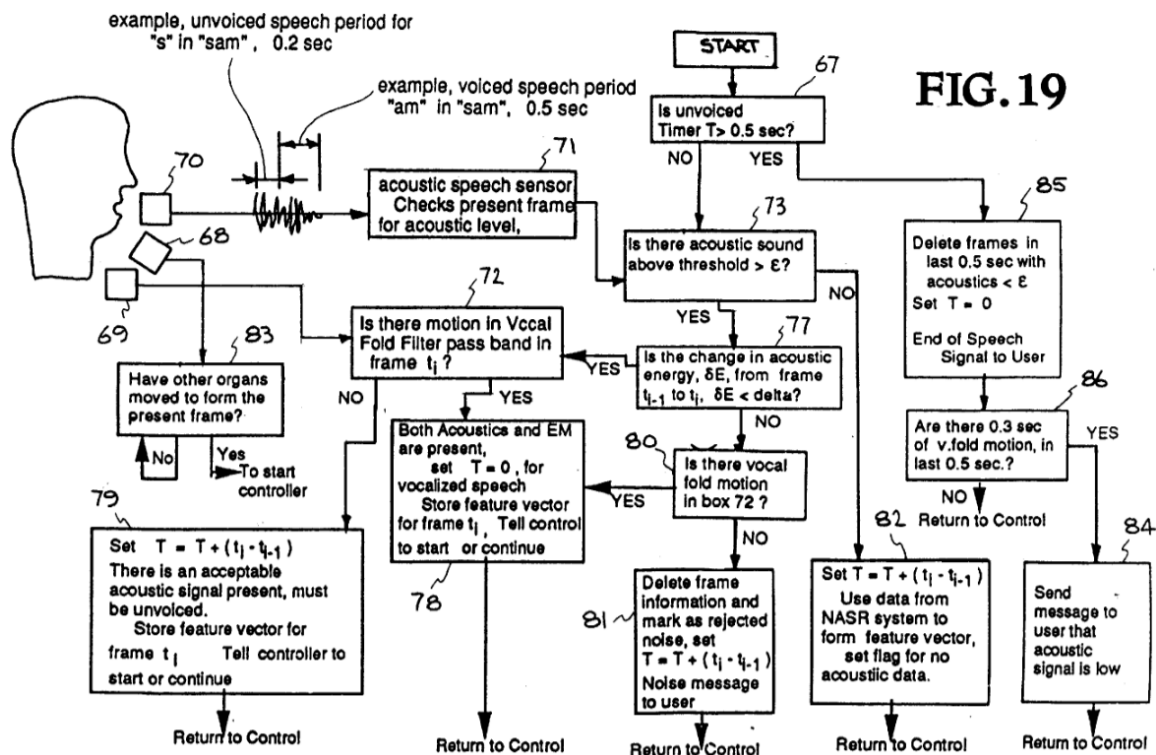


Figure 19 illustrates “a flow chart of an algorithmic procedure for start of speech, end of speech, identification of voiced or unvoiced phoneme, presence of pause, and extraneous noise presence.” Ex. 1003, 15:62–65, 31:9–12, Fig. 19. As Figure 19 shows, nonacoustic information (vocal fold motion) “from EM sensor 69 is input into integrator and band pass filter 72,” and acoustic information “from microphone 70 is inputted into an acoustic speech sensor, A/D converter and 1 sec. memory 71.” *Id.* at 31:15–18, Fig. 19.

In Figure 19, the algorithm begins “at box 67 where a test on timer T for the cumulative duration of unvoiced speech over several speech time frames is conducted (e.g., is $T > 0.5 \text{ sec.}$?).” Ex. 1003, 31:64–67, Fig. 19. “If the duration of testing is shorter than 0.5 sec” and “the acoustic microphone 70 output signal exceeds a preset threshold 73” but does not exceed “a loud noise threshold 77,” in box 72 the “algorithm looks for vocal fold motion to occur within the same” speech time frame t_i . *Id.* at 31:67–32:4, Fig. 19. “If acoustics and vocal fold motion are present, then a voiced speech unit has occurred and the unvoiced timer, T, is set to zero in box 78, and processing on frame t_i continues.” *Id.* at 32:4–7, Fig. 19. “If satisfactory acoustics, but no vocal fold motion occurs in box 72, the sound in frame t_i is labeled as nonvoiced speech in box 79 and processing continues.” *Id.* at 32:7–10, Fig. 19. “If the acoustic signal is too loud in box 77, and vocal fold motion is detected at box 80, the signal is processed as loud voiced speech” in box 78. *Id.* at 32:10–12, Fig. 19. “If acoustic speech is below a threshold box 73 (a test not used for silent or whispered speech applications), the EM Sensor feature vector is processed in box 82, the acoustic Signal is labeled as too weak in box 82.” *Id.* at 32:12–18, Fig. 19.

In another embodiment, Holzrichter discloses background noise suppression using either (1) noise elimination by an onset of speech algorithm or (2) elimination of acoustic noise that enters a microphone during speech by using non-acoustic data comparisons. Ex. 1003, 33:19–28.

2. Overview of Takano (Exhibit 1005)

Takano is a Japanese unexamined patent application publication titled “Sound Collection Device, Voice Recognition Device, Methods Thereof,

and Program Recording Medium,” filed on August 19, 1998, and published on November 5, 1999. Ex. 1005, codes (11), (12), (22), (43), (54). Takano states that the “invention relates to a sound collection device for suppressing ambient noise components in a sound source signal containing ambient noise and extracting a target signal, a voice recognition device which has the sound collection function of the sound collection device and achieves a high recognition rate, methods thereof,” and “a program recording medium.” *Id.* ¶ 1; *see id.* ¶¶ 8–18, code (57).

Takano describes problems estimating noise in one-microphone systems and in two-microphone systems, e.g., when the two microphones are either too close or too far apart. Ex. 1005 ¶¶ 4–5. Takano discloses several embodiments of systems that overcome problems estimating noise and “improve sound quality.” *See, e.g., id.* ¶¶ 8–31, 39–42, 50–54, 59, Figs. 1–3, 6A–6C, 7–9, 14–15.

Takano’s Figure 2 (reproduced below) depicts an embodiment of a two-microphone system that overcomes problems estimating noise:

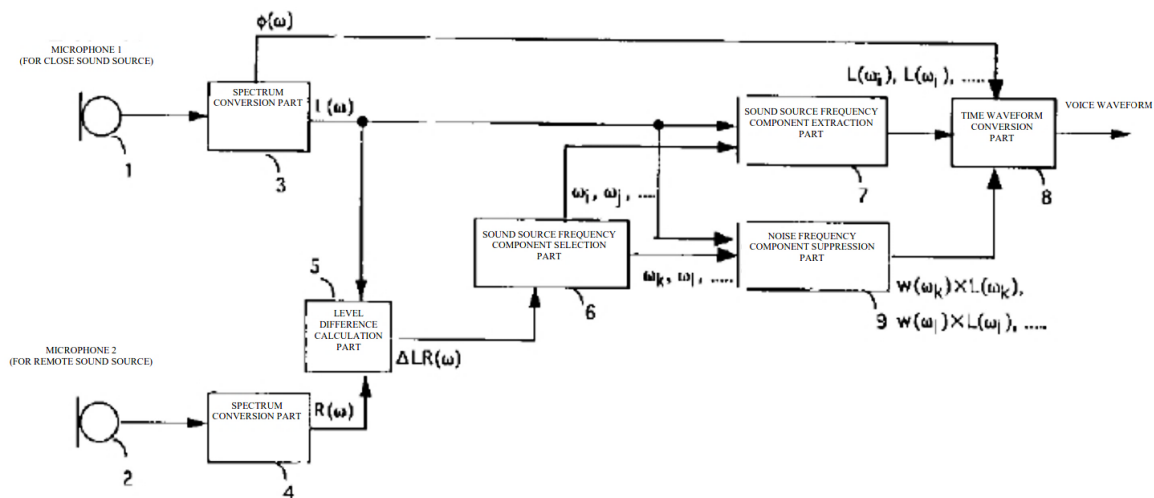


FIG. 2

Figure 2 illustrates a two-microphone system including the following components:

- (1) microphone 1 close to a sound source and producing signal L;
- (2) microphone 2 more remote from the sound source and producing signal R;
- (3) spectrum conversion part 3 receiving an input from microphone 1;
- (4) spectrum conversion part 4 receiving an input from microphone 2;
- (5) level difference calculation part 5 receiving inputs from spectrum conversion part 3 and spectrum conversion part 4;
- (6) sound source frequency component selection part 6;
- (7) sound source frequency component extraction part 7;
- (8) time waveform conversion part 8; and
- (9) noise frequency component suppression part 9.

Ex. 1005 ¶¶ 22–27, Fig. 2. Takano discloses that that “the level difference in the outputs of the microphones 1 and 2 for a sound from the sound source of the target signal is large, but a sound source of noise is sufficiently far from the microphones 1 and 2, and the level difference in the outputs of the microphones 1 and 2 due to noise is zero or minimal.” *Id.* ¶ 22.

Takano's Figure 8 (reproduced below) depicts an operating procedure for the Figure 2's two-microphone system:

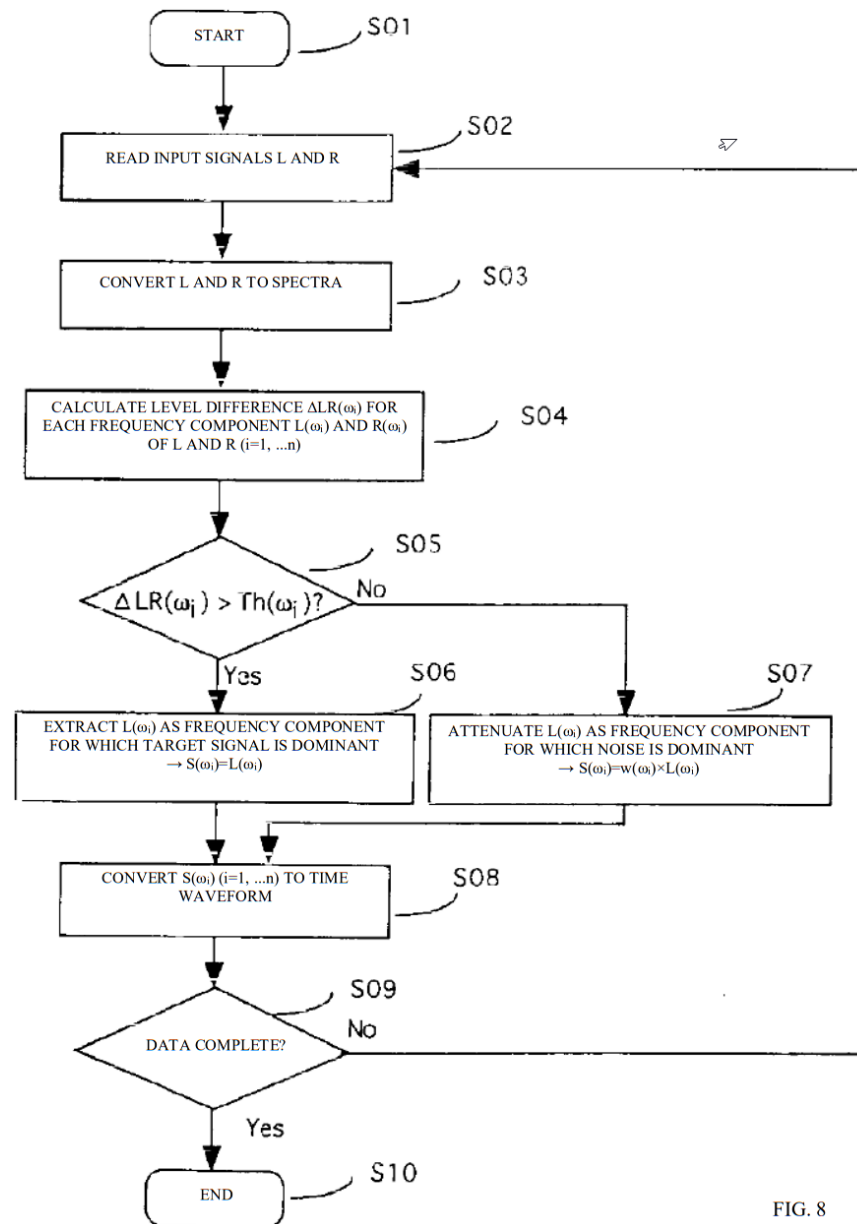


FIG. 8

Figure 8 illustrates a flowchart with steps S01 through S10. Ex. 1005 ¶ 27, Fig. 8.

In step S02, microphone 1 provides signal L as an input to spectrum conversion part 3 and microphone 2 provides signal R as an input to spectrum conversion part 4. Ex. 1005 ¶¶ 23, 27, Figs. 2, 8.

In step S03, spectrum conversion part 3 converts signal L to spectra $L(\omega_i)$ and spectrum conversion part 4 converts signal R to spectra $R(\omega_i)$, where “ (ω_i) is the frequency ($i = 1, 2, \dots, n$).” Ex. 1005 ¶¶ 23–24, Fig. 8.

In step S04, level difference calculation part 5 calculates the level difference $\Delta LR(\omega_i)$ according to the following equation where $|L(\omega_i)|$ is the amplitude of frequency component $L(\omega_i)$ and $|R(\omega_i)|$ is the amplitude of frequency component $R(\omega_i)$:

$$\Delta LR(\omega_i) = 20 \log_{10}(|L(\omega_i)|/|R(\omega_i)|)$$

Ex. 1005 ¶ 24, Fig. 8.

In step S05, sound source frequency component selection part 6 compares the level difference $\Delta LR(\omega_i)$ to “a preset threshold value $Th(\omega_i)$ for each frequency component.” Ex. 1005 ¶ 25, Figs. 2, 8. If $\Delta LR(\omega_i)$ is greater than $Th(\omega_i)$, i.e., $\Delta LR(\omega_i) > Th(\omega_i)$, “the target signal is dominant.” *Id.* ¶ 25. If $\Delta LR(\omega_i)$ is less than or equal to $Th(\omega_i)$, i.e., $\Delta LR(\omega_i) \leq Th(\omega_i)$, “the target signal is not dominant” and instead noise is dominant. *Id.* “The threshold value $Th(\omega_i)$ can be determined experimentally from the correlation between the microphones 1 and 2 and the target signal sound source (for example, the mouth of the user of the microphones 1 and 2).” *Id.* ¶ 26.

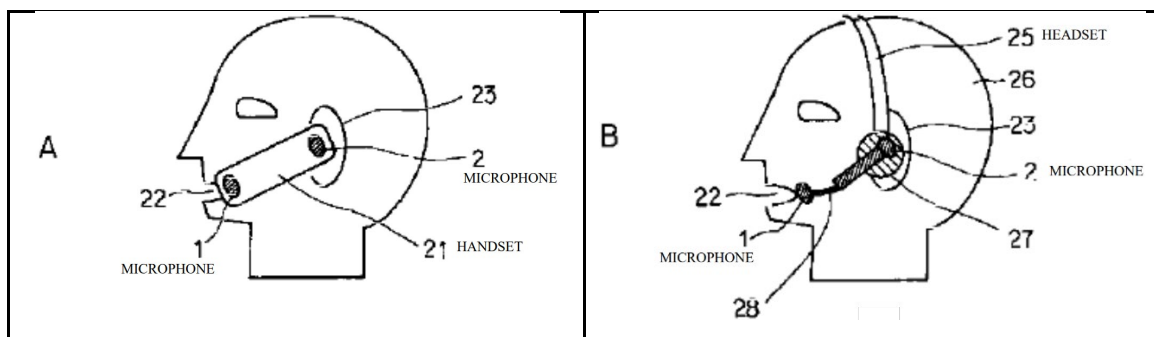
If “the target signal is dominant” ($\Delta LR(\omega_i) > Th(\omega_i)$), in step S06 sound source frequency component extraction part 7 extracts frequency component $L(\omega_i)$ and provides the extracted frequency component to time waveform conversion part 8. Ex. 1005 ¶¶ 25–26, 28, Figs. 2, 8.

But if “the target signal is not dominant” and instead noise is dominant ($\Delta LR(\omega_i) \leq Th(\omega_i)$), in step S07 noise frequency component suppression part 9 applies a weighting factor $w(\omega_i)$ to frequency component

$L(\omega_i)$ to attenuate the frequency component and provides the attenuated frequency component to time waveform conversion part 8. Ex. 1005 ¶¶ 26–28, Figs. 2, 8. The weighting factor $w(\omega_i)$ may have a “small value” (near 0) when “the amount of noise suppression is to be emphasized” and a “large value” (near 1) when “sound quality is to be emphasized.” *Id.* ¶ 28. Applying the weighting factor $w(\omega_i)$ to frequency components where “the target signal is not dominant” and instead noise is dominant addresses the problem of “sound quality degradation due to the dropping of frequency components of the target signal that were present in a frequency band not extracted as a target signal.” *Id.* ¶ 54.

Time waveform conversion part 8 receives the output of sound source frequency component extraction part 7 and the output of noise frequency component suppression part 9. Ex. 1005 ¶ 29, Fig. 2. In step S08, time waveform conversion part 8 applies “an inverse Fourier transform” to the received signals and outputs the transformed signals as “a time waveform.” *Id.* ¶ 29, Fig. 8; *see id.* ¶ 26.

Takano’s Figures 6A and 6B (reproduced below) depict embodiments of two-microphone systems that overcome problems estimating noise:



Above on the left, Figure 6A illustrates microphone 1 and microphone 2 mounted on handset 21 located near a user’s head, with microphone 1 positioned “near the mouth 22 of the user” and microphone 2 positioned

“near the ear 23.” Ex. 1005 ¶ 50, Fig. 6A. Above on the right, Figure 6B illustrates microphone 1 and microphone 2 mounted on headset 25 affixed to a user’s head, with microphone 1 positioned “near the mouth 22” at “the free end of the support arm 28” and microphone 2 attached to “a portion of a receiver 27 in contact with the ear 23.” *Id.* ¶ 51, Fig. 6B.

Additionally, Takano discloses using either (1) a “bone conduction” microphone, (2) an “external auditory meatus” microphone, or (3) a camera as a “voicing detection part” when “determining the voiced/voiceless state of a sound source of a target signal.” Ex. 1005 ¶¶ 19, 39–42, 49, Figs. 14–15. Takano explains that “when a camera is used for the voicing detection part,” the “voiced state determination may be performed by monitoring fluctuation in a video signal of the mouth.” *Id.* ¶ 42; *see id.* ¶ 19.

3. Overview of Kinsler (Exhibit 1004)

Kinsler is a book titled “Fundamentals of Acoustics,” by Lawrence E. Kinsler et al., published no later than 1982 by John Wiley & Sons, Inc. Ex. 1004, i–ii;⁵ *see* Ex. 1019 ¶¶ 5, 22–28. Kinsler explains that the “physical characteristics of speech, music, and noise can be measured with considerable precision by standard acoustic instrumentation including microphones, filters, spectrum analyzers, and oscilloscopes, and the results of such measurements may be expressed in terms of precise physical parameters.” Ex. 1004, 246 (emphases omitted).

⁵ Except for the page numbers i–xi that Petitioner applied to the cover page and the introductory matter in Exhibit 1004 (Kinsler), we follow Petitioner’s practice and cite to the page numbers that appear in the book rather than the page numbers that Petitioner applied to the exhibit.

For detecting a signal in the presence of noise, Kinsler explains that:

- (1) when “attempting to isolate desired information, *signal*, from undesired information, *noise*,” the “detection of a *signal* in the presence of *noise* ultimately reduces to a statistical problem”;
- (2) where τ represents a time interval and A_i represents the output of a detection system for each time $t_i = t_{i-1} + \tau$, “the only way to decide if a signal is present is to select a *threshold criterion* A_T and assume that if $A_i > A_T$ there is a signal present and if $A_i < A_T$ there is no signal”; and
- (3) “choosing a particular criterion A_T amounts to finding a suitable balance between the probability of a detection and the probability of a false alarm.”

Ex. 1004, 252–53 (emphases in original).

Kinsler discloses the following equation for a detection threshold DT where S represents the signal power and N represents the noise power:

$$DT = 10 \log(S/N)$$

Ex. 1004, 256. Kinsler discusses “two extremes” for relating this equation to “the properties observed” at a detection system output A. *Id.* First, “[i]f the signal is known exactly, then the optimum processing system is based on cross correlation, which searches for a facsimile of the known signal in the received signal with noise.” *Id.* (emphasis omitted). Second, “[i]f the properties of the signal are completely known, square-law processing is best,” e.g., a detection system “receives acoustic pressure, squares it to determine the acoustic energy in each observation interval, and presents that as the output A.” *Id.* (emphasis omitted). Kinsler identifies a detection threshold DT for each of these two processing schemes. *Id.* at 257.

Kinsler discloses the following equation for a cross correlation function $F(T)$ between a “known signal,” e.g., human speech, and a

“received signal,” where $s(t)$ represents the “known signal,” $r(t + T)$ represents the “received signal,” T represents an adjustable time in the correlator, and τ represents a time interval:

$$F(T) = \int_{\tau} s(t)r(t + T) dt$$

Ex. 1004, 256. $F(T)$ “plays the role of the detection system output A .”

Id.

4. Overview of Lyons (Exhibit 1008)

Lyons is a book titled “Understanding Digital Signal Processing,” by Richard G. Lyons, published no later than 1997 by Addison-Wesley Publishing Co. Ex. 1008, i–iv;⁶ see Ex. 1019 ¶¶ 8, 41–56. Lyons discloses digital signal-processing techniques for a wide variety of technical fields, including speech recognition. See Ex. 1008, 1, 21, 47, 127–28, 155–56, 217–18, 295, 317–18, 441–42.

Lyons explains that “unwanted fluctuations in a measured value are called *noise*.” Ex. 1008, 319 (emphasis in original). Lyons discloses using “statistical measures known as the mean [or average], variance, and standard deviation” for quantitative analyses. *Id.* at 319–20; see *id.* at 476–84. Lyons explains that if “the sequence $x(n)$ represents a time series of signal samples,” then:

- (1) “ x_{ave} specifies the constant, or DC, value of the signal”;
- (2) “the standard deviation σ reflects the amount of the fluctuating, or AC, component of the signal”; and

⁶ Except for the page numbers i–xvi that Petitioner applied to the cover page and the introductory matter in Exhibit 1008 (Lyons), we follow Petitioner’s practice and cite to the page numbers that appear in the book rather than the page numbers that Petitioner applied to the exhibit.

- (3) “the variance σ^2 is an indication of the power in the fluctuating component.”

Id. at 320.

For a time series of signal samples $x(n)$, Lyons discloses the following equations for the standard deviation σ and the variance σ^2 :

$\sigma = \sqrt{\frac{1}{N} \sum_{n=1}^N [x(n) - x_{\text{ave}}]^2}$	$\sigma^2 = \frac{1}{N} \sum_{n=1}^N [x(n) - x_{\text{ave}}]^2$
--	---

The equation above on the right for the variance σ^2 shows the variance σ^2 calculated as the sum of the squared difference between each signal $x(n)$ and the average x_{ave} divided by the number signals N . Ex. 1008, 320, 478.

The equation above on the left for the standard deviation σ shows the standard deviation σ calculated as the positive square root of the variance σ^2 .

Id. at 320, 479.

5. Overview of Torick (Exhibit 1006)

Torick is a paper titled “An Interphone System for ‘Hands-Free’ Operation in High Ambient Noise,” by E.L. Torick and R.G. Allen, published in the IEEE Transactions on Audio and Electroacoustics, Volume AU-14, Number 4, in December 1966. Ex. 1006, 168;⁷ see Ex. 1019 ¶¶ 6–7, 29–40.

Torick describes the problem that “maximum noise levels produced by military vehicles have increased by several orders of magnitude” over “the past years” and “direct voice communication in many circumstances

⁷ For Exhibit 1006 (Torick), we follow Petitioner’s practice and cite to the page numbers that appear in the publication rather than the page numbers that Petitioner applied to the exhibit.

has become impossible in armored vehicles or in the presence of jet and rotary wing aircraft.” Ex. 1006, 168. To address that problem, Torick discloses a two-microphone system with a noise-cancelling speech microphone to detect speech and a noise-sensing microphone to detect ambient noise. *Id.* at 168–69.

The system uses the two microphones to implement voice-actuated switching (VOX) and “eliminate the press-to-talk switch.” Ex. 1006, 168–69. “The noise-sensing microphone output is amplified and equalized to correspond to the far-field response of the noise-cancelling speech microphone.” *Id.* at 169. The system continuously compares the “outputs of both microphones.” *Id.* When the “outputs of both microphones” are equal, “the switching circuit is inactive.” *Id.* But when “speech sounds are present at the appropriate microphone, the switching circuit energizes.” *Id.* The “method is analogous to the presence of a virtual threshold” that “rises and falls with corresponding changes in noise levels.” *Id.*

Torick’s Figure 4 (reproduced below) depicts a block diagram of a two-microphone system with a noise-cancelling speech microphone to detect speech and a noise-sensing microphone to detect ambient noise:

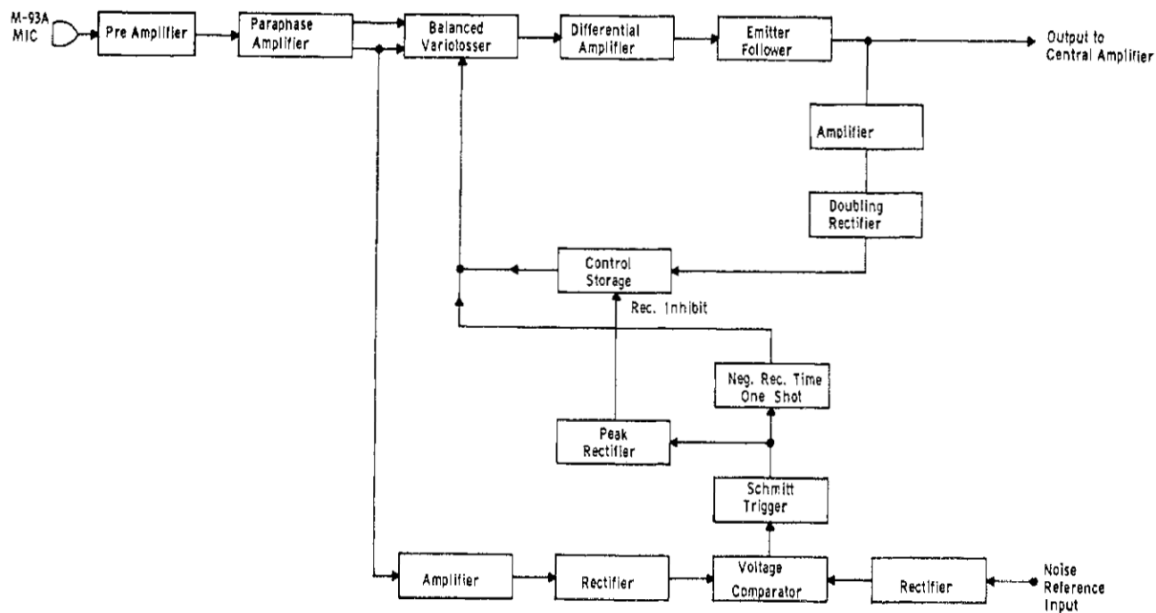


Fig. 4. Crew member's amplifier and control unit.

Figure 4 illustrates an amplifier and control unit “containing VOX circuitry and automatic gain control (AGC)” with a noise-cancelling speech microphone identified as “M-93A MIC” and a noise-sensing microphone identified as “Noise Reference Input.” Ex. 1006, 169–70.

According to Torick, the system provides “highly reliable voice-actuated switching (VOX) at widely varying ambient levels” and permits operation in high ambient noise. Ex. 1006, 168.

6. Overview of Kaneda (Exhibit 1007)

Kaneda is a U.S. patent titled “Method of Detecting Acoustic Signal,” filed on March 8, 1990, and issued on May 4, 1993. Ex. 1007, codes (11), (22), (45), (54). Kaneda states the “invention relates to a method of detecting an acoustic signal, and a method of detecting a period of a desired acoustic signal in a signal including noise and the desired acoustic signal.” *Id.* at 1:5–9; *see id.* at code (57).

Kaneda's Figures 2(a)–2(b) (reproduced below) depict microphone arrangements for detecting speech in the presence of noise where two microphones are located to cause a signal-to-noise ratio (S/N ratio) difference between the outputs from the two microphones:

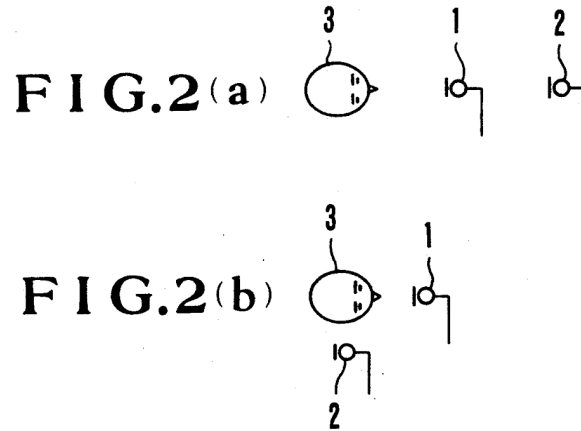


Figure 2(a) (the upper figure above) illustrates “a first microphone 1” located “near a speaker 3” and “a second microphone 2” located “away from the speaker 3.” Ex. 1007, 1:55–58, Fig. 2(a). Figure 2(b) (the lower figure above) illustrates “the first microphone 1” located “in front of the speaker 3” and “the second microphone 2” located “near the side of the speaker 3.” *Id.* at 1:58–61, Fig. 2(b).

In each microphone arrangement, “the speech power level of the output from the first microphone is higher than that from the second microphone.” Ex. 1007, 1:61–63. Additionally, “assuming that noise is generated in a remote location, the noise power levels of the outputs from these microphones are almost equal to each other.” *Id.* at 1:64–66. Hence, “an S/N ratio difference in outputs of the two microphones occurs.” *Id.* at 1:66–68.

Kaneda's Figures 3(a)–3(c) (reproduced below) help explain how the microphone arrangements in Figures 2(a)–2(b) operate:

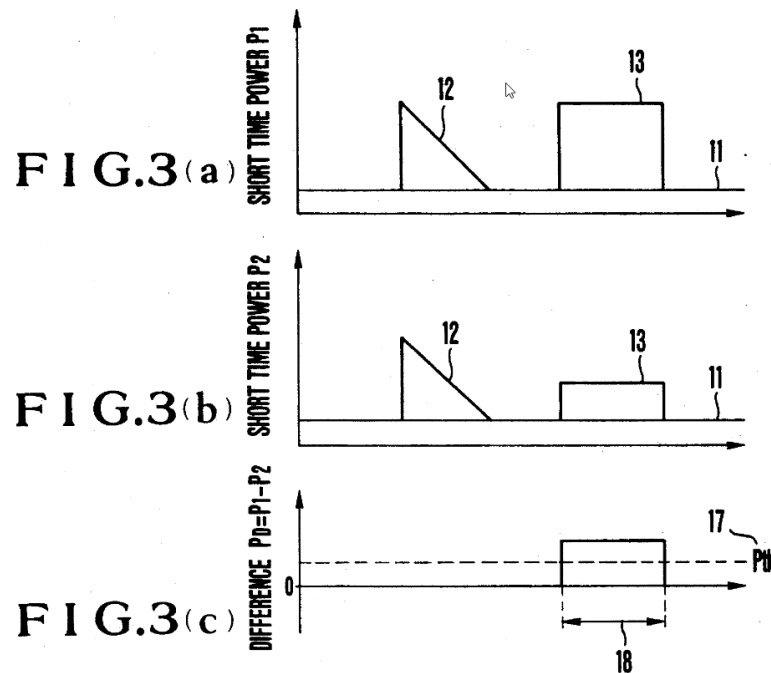


Figure 3(a) (the upper figure above) “shows a time change in power P1 of the output from the first microphone.” Ex. 1007, 2:3–4, Fig. 3(a).

Figure 3(b) (the middle figure above) “shows a time change in power P2 of the output from the second microphone.” *Id.* at 2:5–6, Fig. 3(b). Figure 3(c) (the lower figure above) shows a time change in the power difference PD between the powers P1 and P2 ($PD = P1 - P2$) in Figures 3(a) and 3(b). *Id.* at 2:12–15, Fig. 3(c).

Figures 3(a) and 3(b) each show stationary noise 11, unstationary noise 12, and speech 13. Ex. 1007, 2:6–8, Figs. 3(a)–3(b). With the two microphones arranged as illustrated in Figure 2(a) or 2(b), the noise power levels are the same in Figures 3(a) and 3(b), but the speech power level in Figure 3(b) “is lower than that in” Figure 3(a). *Id.* at 2:8–12, Figs. 3(a)–3(b).

As Figure 3(c) shows, “[w]hen the power difference PD is larger than a given threshold value Pth17, a corresponding time period 18 is detected as a speech period.” Ex. 1007, 2:15–18, Fig. 3(c). Hence, “the unstationary noise period having a high power is not detected as a speech period.” *Id.* at 2:19–21.

*E. Alleged Obviousness over Holzrichter,
Takano, and Kinsler: Claims 1–5*

Petitioner contends that claims 1–5 are unpatentable under § 103(a) as obvious over Holzrichter, Takano, and Kinsler. *See* Pet. 15, 18–43. Patent Owner disputes Petitioner’s contentions. *See* Prelim. Resp. 8–12. Below, we provide overviews of Holzrichter, Takano, and Kinsler, and then we consider the obviousness issues. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Takano, and Kinsler teach or suggest the subject matter in claims 1–5.

1. Analysis of Independent Claim 1

a) Preamble

Claim 1 recites “[a] system for detecting voiced and unvoiced speech in acoustic signals having varying levels of background noise.” Ex. 1001, 11:8–9.

Petitioner contends that Holzrichter teaches claim 1’s preamble because Holzrichter discloses “a system and method for detecting the ‘start of speech, end of speech, identification of voiced or unvoiced phoneme, presence of pause, and extraneous noise presence.’” Pet. 17 (quoting Ex. 1003, 15:63–65) (citing Ex. 1002 ¶¶ 61–62; Ex. 1003, 15:62–65, 22:6–28, 31:8–14, 34:28–67, 52:66–53:2, Figs. 9A–9B, 19).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1's preamble. *See, e.g.*, Prelim. Resp. 6–13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware, LLC v. Nat'l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015).

Generally, a preamble does not limit a claim. *Allen Eng'g Corp. v. Bartell Indus., Inc.*, 299 F.3d 1336, 1346 (Fed. Cir. 2002). We need not decide whether claim 1's preamble limits the claim because Petitioner establishes sufficiently for purposes of institution that Holzrichter teaches claim 1's preamble. *See* Pet. 17; Ex. 1002 ¶¶ 60–62. As Petitioner contends, Holzrichter discloses “a system and method for detecting the ‘start of speech, end of speech, identification of voiced or unvoiced phoneme, presence of pause, and extraneous noise presence.’” Ex. 1003, 15:62–65, Fig. 19; *see id.* at 31:9–12, 52:66–53:2; Ex. 1002 ¶ 61; Pet. 17. For example, Holzrichter explains that Figure 19 shows “one method for determining the onset of speech, noise rejection, voiced-unvoiced speech, end of speech, and pauses using acoustic and non-acoustic information together.” Ex. 1003, 31:9–12; *see* Ex. 1002 ¶ 61.

b) At Least Two Microphones

Claim 1 recites “at least two microphones that receive the acoustic signals.” Ex. 1001, 11:11.

Petitioner contends that the combined disclosures in Holzrichter and Takano teach claim 1's “two microphones” limitation. *See* Pet. 18–22. Specifically, Petitioner asserts that:

- (1) Holzrichter discloses “a microphone (70) for receiving acoustic signals” and “comparing the signal strength

against pre-set thresholds to determine whether the signals are speech”; and

- (2) Takano discloses “a similar process but uses two microphones for receiving acoustic signals” and “as in Holzrichter, us[ing] the acoustic energy of the signals received by the microphones to determine whether the system should further analyze the signals for speech.”

Id. at 18–20.

To support its position, Petitioner provides a side-by-side comparison of Takano’s Figure 6A and the ’058 patent’s Figure 7 as reproduced below (Pet. 20):

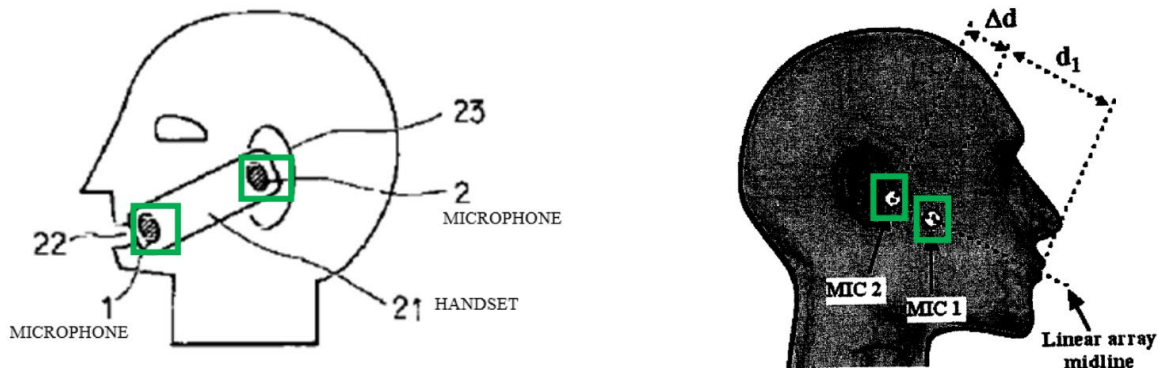


Figure 7

Above on the left, Takano’s Figure 6A illustrates microphone 1 and microphone 2 mounted on handset 21 located near a user’s head, with microphone 1 positioned “near the mouth 22 of the user” and microphone 2 positioned “near the ear 23.” Ex. 1005 ¶ 50, Fig. 6A; *see* Ex. 1002 ¶ 42. The above annotated version of Figure 6A includes a green box around microphone 1 and another green box around microphone 2. *See* Pet. 20; Ex. 1002 ¶ 68. Above on the right, the ’058 patent’s Figure 7 illustrates a user’s head and a two-microphone array with “the microphones Mic 1 and Mic 2 in a linear array with the mouth on the array midline.” Ex. 1001, 2:3–4, 7:52–55, Fig. 7. The above annotated version of Figure 7 includes

a green box around Mic 1 and another green box around Mic 2. *See* Pet. 20; Ex. 1002 ¶ 68.

Petitioner also asserts that “like the ’058 patent, Takano discloses using two microphones to receive acoustic signals and then comparing the difference in the gain of the two microphones against a pre-set threshold to determine whether the signals are speech.” Pet. 22 (citing Ex. 1002 ¶ 72).

At this stage of the proceeding, Patent Owner does not dispute that (1) Holzrichter discloses a one-microphone system that “determine[s] whether there is noise or unvoiced speech” and (2) Takano discloses a two-microphone system.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter and Takano teach claim 1’s “two microphones” limitation. *See* Pet. 18–22; Ex. 1002 ¶¶ 63–72. Specifically, Takano discloses several embodiments of two-microphone systems that overcome problems estimating noise and “improve sound quality.” *See, e.g.*, Ex. 1005 ¶¶ 8–18, 20–31, 50–54, 59, Figs. 1–3, 6A–6C, 7–9.

As an example, Takano’s Figure 6A illustrates microphone 1 and microphone 2 mounted on handset 21 located near a user’s head, with microphone 1 positioned “near the mouth 22 of the user” and microphone 2 positioned “near the ear 23.” Ex. 1005 ¶ 50, Fig. 6A; *see* Ex. 1002 ¶ 42. As another example, Takano’s Figure 6B illustrates microphone 1 and microphone 2 mounted on headset 25 affixed to a user’s head, with microphone 1 positioned “near the mouth 22” at “the free end of the support

arm 28” and microphone 2 attached to “a portion of a receiver 27 in contact with the ear 23.” Ex. 1005 ¶ 51, Fig. 6B.

c) At Least One Voicing Sensor

Claim 1 recites “at least one voicing sensor that receives physiological information associated with human voicing activity.” Ex. 1001, 11:12–13.

Petitioner contends that Holzrichter teaches claim 1’s “voicing sensor” limitation because Holzrichter discloses “a radar sensor (EM sensor 68) that is used to detect voicing activity.” Pet. 25 (citing Ex. 1002 ¶ 80; Ex. 1003, 8:25–49, 9:23–54, 10:50–65, 11:66–12:11, 31:17–18, 32:31–35, Figs. 4–5, 8A–8C, 12–13, 24A–24B).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1’s “voicing sensor” limitation. *See, e.g.*, Prelim. Resp. 6–13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that Holzrichter teaches claim 1’s “voicing sensor” limitation. *See* Pet. 26; Ex. 1002 ¶¶ 79–80.

d) At Least One Processor

Claim 1 recites “at least one processor coupled among the microphones and the voicing sensor.” Ex. 1001, 11:15–16.

Petitioner contends that Holzrichter teaches claim 1’s “processor” limitation because Holzrichter discloses “a processor coupled to microphone 70 and EM sensors 68.” Pet. 27 (citing Ex. 1002 ¶ 82; Ex. 1003, 9:55–10:14, 18:24–67, 31:15–20, Figs. 5, 13). Also, Petitioner asserts that “when adding the second microphone of Takano to Holzrichter’s

system,” an ordinarily skilled artisan would have coupled the second microphone “to the processor to allow Holzrichter’s system to use the second microphone to identify noise and speech.” *Id.* (citing Ex. 1002 ¶ 82).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1’s “processor” limitation. *See, e.g.*, Prelim. Resp. 6–13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that Holzrichter teaches claim 1’s “processor” limitation. *See* Pet. 26; Ex. 1002 ¶¶ 81–83.

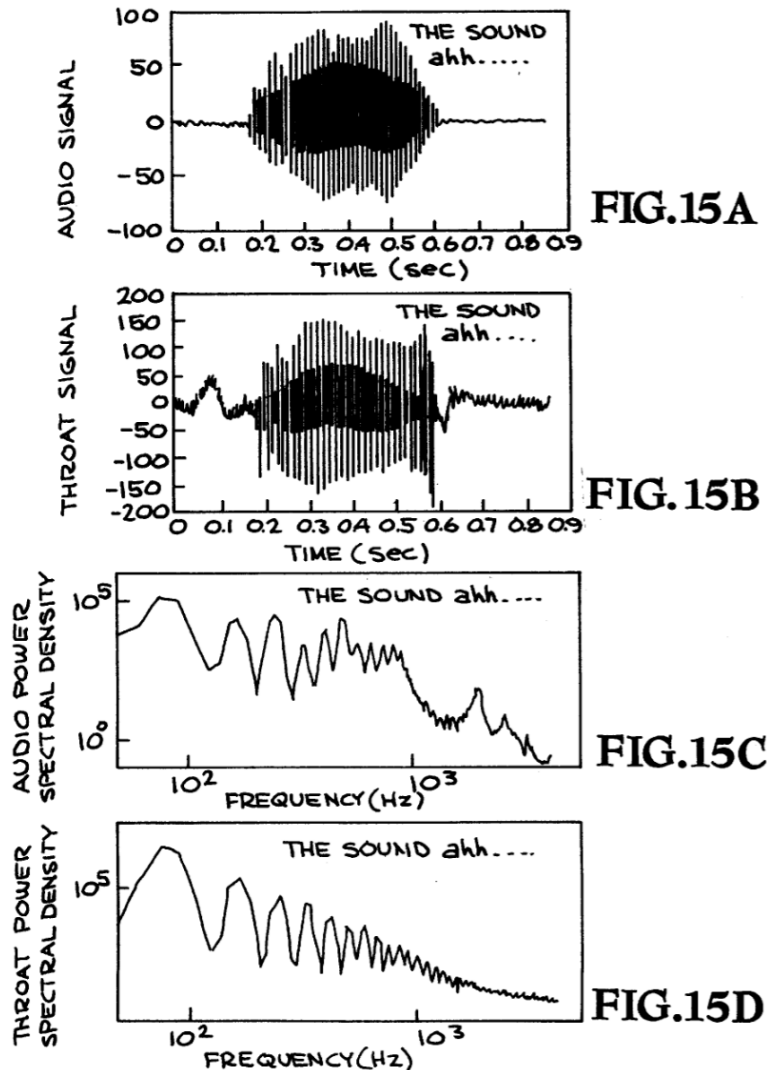
e) Processor Generates Cross Correlation Data

Claim 1 specifies that the processor “generates cross correlation data between the physiological information and an acoustic signal received at one of the two microphones.” Ex. 1001, 11:18–20.

Petitioner contends that “Holzrichter, alone or in combination with Kinsler,” renders obvious the “cross correlation data” limitation. *See* Pet. 27–30. Regarding Holzrichter alone, Petitioner asserts that Holzrichter’s system “determines whether both speech and vocal fold motion are present within a certain speech frame based on the physiological information received at EM sensor 69 and an acoustic signal received at microphone 70.” *Id.* at 28 (citing Ex. 1002 ¶ 85; Ex. 1003, 31:61–32:27, Figs. 15A–15D, 19). Petitioner also asserts that “Holzrichter further discloses that the signal it receives from the EM throat sensor and the audio signal received by the microphone are closely correlated.” *Id.* at 29–30

(citing Ex. 1002 ¶ 87; Ex. 1003, 17:4–23, 20:11–34, 26:44–27:4, 31:61–32:27, Figs. 15A–15D).

To support its assertion about the correlation between the EM-sensor signal and the audio (acoustic) signal, Petitioner provides Holzrichter's Figures 15A–15D as reproduced below (Pet. 30):



Figures 15A–15D illustrate EM-sensor data and audio (acoustic) data for the phoneme “/ah/” as “amplitude vs. time and with the Fourier power spectrum.” Ex. 1003, 15:49–51, 26:44–45, Figs. 15A–15D. According to Petitioner, Figures 15A–15D exemplify “how easily” speech organ location

information “can be correlated with simultaneous acoustic word signals from a microphone.” Pet. 29 (quoting Ex. 1003, 26:57–60).

Petitioner acknowledges that “Holzrichter does not expressly state that the signals can be correlated using a cross correlation function.” Pet. 30. But Petitioner asserts that “using such a function would have been obvious.” *Id.* (citing Ex. 1002 ¶ 88).

As evidence that “using such a function would have been obvious,” Petitioner cites Kinsler. Pet. 29–30. Petitioner contends that Kinsler discloses:

- (1) the “optimum processing system” to search for a “known signal” (e.g., the audio signal from a microphone) in a “received signal” (e.g., the vocal movement signal from the EM sensor) is “cross correlation”;
- (2) calculating “cross correlation data” using a cross correlation function $F(T)$; and
- (3) comparing the result of the cross correlation function $F(T)$ to a detection threshold DT to determine whether “the two signals are correlated,” i.e., to determine whether the result of the cross correlation function $F(T)$ exceeds the detection threshold DT .

Id. at 31 (citing Ex. 1002 ¶ 90; Ex. 1004, 252, 256–57).

Based on the current record, including Dr. Stern’s supporting testimony, we are persuaded Petitioner establishes sufficiently for purposes of institution that Holzrichter teaches the “cross correlation data” limitation and that the combined disclosures in Holzrichter and Kinsler also teach the limitation. *See* Pet. 27–31; Ex. 1002 ¶¶ 84–90, 92. Specifically, Holzrichter discloses correlating EM-sensor data and audio (acoustic) data. Ex. 1003, 15:49–51, 16:27–29, 20:11–15, 26:44–27:4, 45:2–15, Figs. 15A–15D, 25A–25H; *see* Ex. 1002 ¶¶ 85, 87, 92.

As an example, Holzrichter explains that “[a]fter an EM pulse (or pulse train) is received it must be processed, correlated with other pulses from other organs and correlated with acoustic data and fed to an algorithm.” Ex. 1003, 20:11–13; *see id.* at 26:57–60; Ex. 1002 ¶¶ 40, 61. As another example, for Figures 15A–15D illustrating EM-sensor data and audio (acoustic) data for the phoneme “/ah/,” Holzrichter explains that “[t]his simple experiment also indicates how easily this speech organ location information can be correlated with simultaneous acoustic word signals from a microphone.” Ex. 1003, 15:49–51, 26:44–45, 26:57–60; *see* Ex. 1002 ¶¶ 87, 92. In yet another example, found in Figures 25A–25H illustrating EM-sensor data and audio (acoustic) data for the sentence “the quick brown fox jumped over the lazy dog’s back,” Holzrichter explains that the figures show “simultaneous acoustic and vocal fold stops” as well as “word starts and stops.” Ex. 1003, 16:27–29, 45:2–12.

Holzrichter describes one process for correlating EM-sensor data and audio (acoustic) data as follows:

The acoustic signals from a microphone sensing the same words as the EM sensor(s) were displayed on a separate trace of the oscilloscope and were sampled in an A/D converter and were stored in a memory with identical start times and with identical time bin numbers as the memory bins for the EM sensor data. The correlation between these two (or more) signals were used to validate the assignment of irreducible spoken sound units (i.e. [phone-like units] PLUs) with the expected location or activity of the vocal organ being measured.

Ex. 1003, 26:60–27:2; *see* Ex. 1002 ¶ 92. Hence, we are persuaded at this stage of the proceeding that Holzrichter discloses correlating EM-sensor data and audio (acoustic) data to determine how physiological information relates

to “spoken sound units” producing acoustic signals received at a microphone.

Thus, Holzrichter teaches “generat[ing] cross correlation data between the physiological information and an acoustic signal received at one of the two microphones” according to claim 1. *See* Ex. 1002 ¶¶ 85, 87, 92; Ex. 1003, 15:49–51, 16:27–29, 20:11–15, 26:44–27:4, 45:2–15, Figs. 15A–15D, 25A–25H. Holzrichter does not disclose a particular algorithm or equation for correlating EM-sensor data and audio (acoustic) data. But the “cross correlation data” limitation does not require a particular algorithm or equation. *See* Ex. 1001, 11:18–20.

In any event, Kinsler discloses the following equation for a cross correlation function $F(T)$ between a “known signal,” e.g., human speech, and a “received signal,” where $s(t)$ represents the “known signal,” $r(t + T)$ represents the “received signal,” T represents an adjustable time in the correlator, and τ represents a time interval:

$$F(T) = \int_{\tau} s(t)r(t + T) dt$$

Ex. 1004, 256; *see* Ex. 1002 ¶¶ 41, 90. $F(T)$ “plays the role of the detection system output A .” Ex. 1004, 256; *see* Ex. 1002 ¶ 90.

Additionally, Kinsler explains that where τ represents a time interval and A_i represents the output of a detection system for each time $t_i = t_{i-1} + \tau$, “the only way to decide if a signal is present is to select a *threshold criterion* A_T and assume that if $A_i > A_T$ there is a signal present and if $A_i < A_T$ there is no signal.” Ex. 1004, 252 (emphasis in original); *see* Ex. 1002 ¶¶ 41, 89.

Dr. Stern testifies that Kinsler’s cross correlation function $F(T)$ may use (1) “the audio signal from a microphone,” e.g., human speech, as the “known signal” and (2) “the vocal movement signal from the EM sensor” as

the “received signal” to cross correlate EM-sensor data and audio (acoustic) data. Ex. 1002 ¶ 90; *see id.* ¶ 41. Dr. Stern also testifies that to “use the cross-correlation method,” a “value from the cross correlation function $F(T)$ would be calculated” and compared to a detection threshold to determine if the calculated value “meets or exceeds the threshold” such that “the two signals are correlated (i.e., one signal is present in the other).” *Id.* ¶ 90 (citing Ex. 1004, 252, 256–57).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1’s “cross correlation” limitation. *See, e.g.*, Prelim. Resp. 6–13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that Holzrichter teaches claim 1’s limitation that the processor “generates cross correlation data between the physiological information and an acoustic signal received at one of the two microphones.” *See* Pet. 27–31; Ex. 1002 ¶¶ 84–90.

f) Processor Identifies Information of the Acoustic Signals as Voiced Speech

Claim 1 specifies that the processor “identifies information of the acoustic signals as voiced speech when the cross correlation data corresponding to a portion of the acoustic signal received at the one receiver exceeds a correlation threshold.” Ex. 1001, 11:21–24.

Petitioner contends that the combined disclosures in Holzrichter and Kinsler teach claim 1’s “identifies voiced speech” limitation. *See* Pet. 34–35. Specifically, Petitioner asserts that:

- (1) Holzrichter's system "identifies information in the acoustic signal as voiced speech when it determines that the vocal motion signal is correlated to the acoustic signal"; and
- (2) "Kinsler discloses that the optimum way to determine whether two signals are correlated is to calculate a value (cross correlation data) using a cross correlation function $F(T)$ and then compare it against a threshold DT " such that "the two signals are correlated (i.e., one signal is present in the other)" if the calculated value exceeds the threshold.

Id. at 33–34 (citing Ex. 1002 ¶ 98; Ex. 1003, 31:61–32:27, Figs. 15A–15D, 19; Ex. 1004, 252, 256–57).

According to Petitioner, the combined disclosures in Holzrichter and Kinsler teach "identifying information in the acoustic signal as voiced speech when the cross correlation data (the value of $F(t)$) exceeds a threshold (DT).” Pet. 35 (citing Ex. 1002 ¶ 99).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1's "identifies voiced speech" limitation. *See, e.g.*, Prelim. Resp. 6–13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern's testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter and Kinsler teach claim 1's "identifies voiced speech" limitation. *See* Pet. 33–35; Ex. 1002 ¶¶ 97–99.

g) Processor Generates Difference Parameters

Claim 1 specifies that the processor "generates difference parameters between the acoustic signals received at each of the two receivers, wherein

the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals.” Ex. 1001, 11:25–29.

Petitioner contends that the combined disclosures in Holzrichter and Takano teach claim 1’s “difference parameters” limitation. *See* Pet. 35–36. Specifically, Petitioner asserts that Takano discloses calculating a “level difference” (difference parameters) for “each frequency component of the signals it receives at its two microphones (portions of the received acoustic signals).” *Id.* at 35 (citing Ex. 1002 ¶ 101; Ex. 1005 ¶¶ 9, 14–15, 23–24, 29, 40–42, 49–51). Petitioner also asserts that Takano discloses calculating a “level difference” based on “the difference in ‘amplitude spectrum components’ (i.e., gain) of the two signals.” *Id.* (quoting Ex. 1005 ¶ 24).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1’s “difference parameters” limitation. *See, e.g.*, Prelim. Resp. 6–13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter and Takano teach claim 1’s “difference parameters” limitation. *See* Pet. 35–36; Ex. 1002 ¶¶ 100–102.

*h) Processor Identifies Information of the
Acoustic Signals as Unvoiced Speech*

Claim 1 specifies that the processor “identifies information of the acoustic signals as unvoiced speech when the difference parameters exceed a gain threshold.” Ex. 1001, 11:30–32.

Petitioner contends that the combined disclosures in Holzrichter and Takano teach claim 1's "identifies unvoiced speech" limitation. *See* Pet. 36–38. Specifically, Petitioner asserts that Holzrichter's system "identifies information of the acoustic signals as unvoiced speech when the strength of the received acoustic signal 'exceeds a preset threshold 73' and no vocal motion is present." *Id.* at 36 (quoting Ex. 1003, 32:1–2) (citing Ex. 1002 ¶ 104; Ex. 1003, 31:61–32:27). Petitioner asserts that Takano discloses:

- (1) comparing the "level difference $\Delta LR(\omega_i)$ " (difference parameter) against a threshold " $Th(\omega_i)$ "; and
- (2) determining that "the target signal is dominant" if the level difference $\Delta LR(\omega_i)$ exceeds the threshold $Th(\omega_i)$, "meaning that the signal is identified as speech."

Id. at 38 (citing Ex. 1002 ¶ 105; Ex. 1005 ¶¶ 9, 14–15, 23–25, 29, 40–42, 49–51, Figs. 1, 6–7, 16–17, 18A).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1's "identifies unvoiced speech" limitation. *See, e.g.,* Prelim. Resp. 6–13. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern's testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter and Takano teach claim 1's "identifies unvoiced speech" limitation. *See* Pet. 36–38; Ex. 1002 ¶¶ 103–105.

*i) Processor Identifies Information
of the Acoustic Signals as Noise*

Claim 1 specifies that the processor “identifies information of the acoustic signals as noise when the difference parameters are less than the gain threshold.” Ex. 1001, 11:33–35.

Petitioner contends that the combined disclosures in Holzrichter and Takano teach claim 1’s “identifies noise” limitation. *See* Pet. 39–40. Specifically, Petitioner asserts that “Takano discloses that, when the level difference is lower than the threshold, ‘the target signal is not dominant,’ meaning that the signal is identified as noise instead of speech.” *Id.* (quoting Ex. 1005 ¶ 25) (citing Ex. 1002 ¶ 108; Ex. 1005 ¶¶ 14–15, 22, 25, 29, 40–42, 49–51, Figs. 1, 6–7, 16–17, 18A).

Patent Owner contests Petitioner’s position and contends that “Petitioner conflates the **absence** of a speech signal being identified as ‘dominant’ with an identification of noise.” Prelim. Resp. 8–9. Patent Owner argues that Takano’s determination that a target signal is “dominant” is a determination that speech is present, or that speech is louder than ambient noise. *Id.* at 10 (citing Ex. 1005 ¶¶ 24–26, 15). But, according to Patent Owner, Takano’s determination that the target signal is “not dominant” does not constitute an identification of acoustic signals because it could just as easily be the presence of silence. *Id.* Rather, Patent Owner contends that Takano does **not** disclose that a “signal is identified as noise instead of speech” when the target signal is not dominant. *Id.* at 9–10.

Patent Owner then argues that the absence of an identification fails to “provide the advantages of the patented invention, which allows calculation

of noise suppression functions during periods where the system is confident that only noise is occurring.” *Id.* at 9.

At this stage and based on the current record, we are persuaded Petitioner has shown sufficiently that the combined teachings of Holzrichter and Takano teach the challenged limitation. Although we understand Patent Owner’s argument that Takano does not explicitly state that it identifying “noise” when a target signal is not domination, Takano does explicitly teach the following:

The level difference in the outputs of the microphones 1 and 2 for a sound from the sound source of the target signal is large, but a sound source of noise is sufficiently far from the microphones 1 and 2, and the level difference in the outputs of the microphones 1 and 2 due to noise is zero or minimal.

Ex. 1005 ¶ 22. Takano further teaches that when the target signal is not dominant, all the frequency components are removed completely so as to improve sound quality and prevent distortion because Takano appears to consider detected frequencies to be noise when the target signal is not dominant. *Id.* ¶¶ 25–26, 29.

Additionally, we preliminarily determine that claim 1 sets the boundary for what is considered “noise.” Specifically, claim 1 recites that if the “difference parameters are less than the gain threshold” then the system “identifies information of the acoustic signal as noise.” *See* Ex. 1001, 11:33–35. Therefore, anything that is less than the gain threshold is considered “noise” by the ’058 patent even if its silence.

Accordingly, based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in

Holzrichter and Takano teach claim 1's "identifies noise" limitation. *See* Pet. 39–40; Ex. 1002 ¶¶ 107–108.

j) Alleged Reasons for Combining the Teachings of the References

Petitioner identifies reasons that would have prompted an ordinarily skilled artisan to combine the teachings of Holzrichter, Takano, and Kinsler in the way Petitioner proposes. *See* Pet. 18, 24–26, 30–31. Additionally, Petitioner explains that an ordinarily skilled artisan "would have had a reasonable expectation of success" in combining the teachings of the references. *Id.* at 26, 33–34. Dr. Stern's testimony supports Petitioner's positions. *See* Ex. 1002 ¶¶ 64, 73–78, 91–96.

For combining the teachings of Holzrichter and Takano, Dr. Stern cites Holzrichter's disclosure that "the strength of acoustic signals (i.e. acoustic energy) could be used to determine whether a signal is speech or noise." Ex. 1002 ¶ 73 (citing Ex. 1003, 31:67–32:10, Fig. 19). He explains that an ordinarily skilled artisan would have been motivated by that disclosure "to look for improved methods for classifying signals as noise or speech based on signal strength, such as the system of Takano." *Id.*

Dr. Stern testifies that adding another microphone to Holzrichter's single-microphone system according to Takano's teachings would advantageously:

- (1) improve Holzrichter's "methods for classifying input signals as either noise or speech based on signal strength" to better distinguish between "a user's speech and noise";
- (2) achieve "a high recognition rate" for a target signal in a signal "containing ambient noise";
- (3) conserve "computing resources" by permitting the system to "perform additional processing only on audio samples that are more likely to be speech"; and

- (4) distinguish between speech and noise more accurately “at high frequencies,” thereby “leading to better sound quality in the processed voice signal.”

Ex. 1002 ¶¶ 64, 74–75 (quoting Ex. 1005 ¶¶ 1, 59). Dr. Stern also testifies that “the proposed combination represents nothing more than a predictable use of prior art elements according to established functions (i.e. using two microphones to detect speech).” *Id.* ¶ 76.

According to Dr. Stern, “Takano’s second microphone would provide an improved method for something Holzrichter already contemplates, namely using the output strength of microphones to distinguish between noise and speech.” Ex. 1002 ¶ 77.

For combining the teachings of Holzrichter and Kinsler, Dr. Stern cites Holzrichter’s disclosure that “the vocal movements detected by Holzrichter’s EM sensor are highly correlated with speech.” Ex. 1002 ¶ 92 (citing Ex. 1003, 17:4–23, 20:11–34, 26:44–27:4, 31:61–32:27). He explains that an ordinarily skilled artisan would have been motivated by that disclosure “to look for other references, such as Kinsler, for methods of determining when two signals are correlated with each other” so that Holzrichter’s system can “detect when the signal received by Holzrichter’s microphone is voiced speech.” *Id.*

Dr. Stern testifies that “Kinsler’s methods would have been a more accurate way of determining that two signals were correlated than Holzrichter’s determination that the audio and vocal movement signals occurred at the same time.” Ex. 1002 ¶ 93. He explains that “using Kinsler’s methods would prevent Holzrichter’s [sic] system from identifying sound from Holzrichter’s microphone that occurred at the same time as some non-voicing vocal movements (such as swallowing) as voice speech.” *Id.*

Dr. Stern also testifies that an ordinarily skilled artisan “would have used Kinsler’s teachings to improve Holzrichter’s method of correlating signals.” *Id.* According to Dr. Stern, “the addition of Kinsler’s methods to Holzrichter would simply improve what Holzrichter already contemplates—correlating an acoustic signal and a vocal movement signal.” *Id.* ¶ 95.

Based on the current record, Petitioner establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had reasons to combine the teachings of Holzrichter, Takano, and Kinsler in the way Petitioner proposes. *See* Pet. 18, 23–25, 30–32; Ex. 1002 ¶¶ 64, 73–77, 91–95. For combining the teachings of Holzrichter and Takano, we credit Dr. Stern’s testimony that adding another microphone to Holzrichter’s single-microphone system according to Takano’s teachings would advantageously improve Holzrichter’s “methods for classifying input signals as either noise or speech based on signal strength” to better distinguish between “a user’s speech and noise.” Ex. 1002 ¶¶ 64, 73–74, 77; *see* Pet. 18, 23.

Based on the current record, we are persuaded that improving Holzrichter’s “methods for classifying input signals as either noise or speech based on signal strength” to better distinguish between “a user’s speech and noise” suffices as a reason for combining the teachings of Holzrichter and Takano. *See* Ex. 1002 ¶¶ 64, 73, 77. 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” references. *KSR*, 550 U.S. at 420.

Dr. Stern identifies an improvement to Holzrichter’s “complete” single-microphone system produced by adding another microphone according to Takano’s teachings, e.g., improved “methods for classifying

input signals as either noise or speech based on signal strength.” Ex. 1002 ¶¶ 64, 73, 77. The “desire to enhance commercial opportunities by improving a product or process is universal” *DyStar Textilfarben GmbH v. C.H. Patrick Co.*, 464 F.3d 1356, 1368 (Fed. Cir. 2006). Hence, improving a “complete” system can provide a reason for combining references. *See Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1380 (Fed. Cir. 2023).

For combining the teachings of Holzrichter and Kinsler, Dr. Stern testifies that “Kinsler’s methods would have been a more accurate way of determining that two signals were correlated than Holzrichter’s determination that the audio and vocal movement signals occurred at the same time.” Ex. 1002 ¶ 93; *see* Pet. 31. At this stage of the proceeding, Patent Owner introduces no contradictory testimony rebutting Dr. Stern. *See* Prelim. Resp. 9–13. Based on the current record, we are persuaded that more accurately “determining that two signals were correlated” suffices as a reason for combining the teachings of Holzrichter and Kinsler. *See* Ex. 1002 ¶ 93. “There is a motivation to combine when a known 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 ‘prior art elements according to their established functions.’” *Intel*, 61 F.4th at 1380–81 (citations omitted) (quoting *KSR*, 550 U.S. at 417; *Intel Corp. v. Qualcomm Inc.*, 21 F.4th 784, 799–800 (Fed. Cir. 2021)).

k) Preliminary Conclusion for Obviousness/Nonobviousness of Claim 1

As discussed above, Petitioner’s analysis explains how the combined disclosures in Holzrichter, Takano, and Kinsler teach claim 1’s subject

matter. *See supra* §§ III.E.1(a)–(j). Additionally, Petitioner provides a reason with rational underpinning as to why an ordinarily skilled artisan would have had a reason to combine the teachings of Holzrichter, Takano, and Kinsler in the way Petitioner proposes and would have had a reasonable expectation of success. *See supra* § III.D.4(j). Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claim 1 unpatentable under § 103(a) as obvious over Holzrichter, Takano, and Kinsler.

2. Analysis of Independent Claim 2

Independent claim 2 recites “[a] method for removing noise from acoustic signals” comprising steps that parallel the processor operations specified in claim 1. Ex. 1001, 12:1–24.

Petitioner asserts that claim 2 is “similar to claim 1” and obvious over Holzrichter, Takano, and Kinsler “for the same reasons as claim 1.” Pet. 40. Dr. Stern’s testimony supports Petitioner’s assertions. *See* Ex. 1002 ¶ 109.

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 2. *See, e.g.*, Prelim. Resp. 8–12. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

For the reasons discussed above, we consider Patent Owner’s argument regarding claim 1 unavailing. *See, e.g., supra* §§ III.E.1(i). Hence, for the reasons discussed above for claim 1, Petitioner demonstrates a reasonable likelihood of prevailing in proving claim 2 unpatentable under § 103(a) as obvious over Holzrichter, Takano, and Kinsler. *See supra* §§ III.E.1(a)–(j).

3. Analysis of Dependent Claims 3–5

Claims 3–5 depend directly from claim 2. Ex. 1001, 12:25–33.

For claims 3–5, Petitioner provides an analysis supported by Dr. Stern’s testimony about how the combined disclosures in Holzrichter and Takano teach the limitation in claim 3 and Holzrichter teaches the limitations in claims 4 and 5. *See* Pet. 40–43; Ex. 1002 ¶¶ 110–115.

At this stage of the proceeding, Patent Owner makes no arguments specific to claims 3–5. *See, e.g.*, Prelim. Resp. 8–12. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the references teach the limitations in claims 3–5. *See* Pet. 40–43; Ex. 1002 ¶¶ 110–115. Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claims 3–5 unpatentable under § 103(a) as obvious over Holzrichter, Takano, and Kinsler.

*E. Alleged Obviousness in view of Holzrichter,
Takano, Kinsler, and Lyons: Claim 3*

Petitioner contends that claim 3 is unpatentable under § 103(a) as obvious over Holzrichter, Takano, Kinsler, and Lyons. *See* Pet. 15, 43–46. Patent Owner disputes Petitioner’s contentions. *See* Prelim. Resp. 12. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Takano, Kinsler, and Lyons teach claim 3’s subject matter.

1. Analysis of Dependent Claim 3

Claim 3 depends directly from claim 2 and further requires “generating the gain threshold using standard deviations corresponding to the generation of the difference parameters.” Ex. 1001, 12:25–27.

Petitioner contends that Takano discloses determining the threshold value $Th(\omega_i)$ for the level difference $\Delta LR(\omega_i)$ empirically. Pet. 43 (citing Ex. 1002 ¶ 118; Ex. 1005 ¶ 26). Petitioner contends that it would have been obvious to an ordinarily skilled artisan in view of Lyons to:

- (1) measure Takano’s level difference $\Delta LR(\omega_i)$ in “various circumstances where Takano’s microphones are receiving noise”;
- (2) “calculate the standard deviation (as taught by Lyons) of the different level differences, which would represent an upper bound for how high Takano’s level difference could reach when receiving noise”; and
- (3) “use this upper bound as the threshold for Takano to determine when it is receiving noise instead of speech.”

Id. at 44 (citing Ex. 1002 ¶ 121).

Patent Owner disputes that Holzrichter, Takano, Kinsler, and Lyons teach claim 3’s subject matter. Prelim. Resp. 12. Specifically, Patent Owner asserts that Petitioner’s arguments for claim 3 “do not remedy the deficiencies” of the challenge to the independent claims. *Id.*

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Takano, Kinsler, and Lyons teach claim 3’s subject matter. *See* Pet. 43–44; Ex. 1002 ¶¶ 117–121.

Based on the current record and for the reasons discussed above for claims 1 and 2, we disagree with Patent Owner that there are deficiencies in the references as applied to claim 3. *See* Prelim. Resp. 12; *supra* §§ III.E.1(a)–(j).

2. Alleged Reasons for Combining the Teachings of the References

Petitioner identifies reasons that would have prompted an ordinarily skilled artisan to combine Lyons’s teachings with the teachings of Holzrichter, Takano, and Kinsler. *See* Pet. 44–46. Additionally, Petitioner explains that an ordinarily skilled artisan “would have had a reasonable expectation of success” in combining the teachings of the references. *Id.* at 46. Dr. Stern’s testimony supports Petitioner’s positions. *See* Ex. 1002 ¶¶ 122–127.

For instance, Dr. Stern testifies that an ordinarily skilled artisan “would have understood, based on Lyons, that calculating the gain threshold using standard deviations corresponding to generation of the difference parameters would ensure that the gain threshold is meaningful.” Ex. 1002 ¶ 123. Dr. Stern also testifies that an ordinarily skilled artisan “would have been motivated to use, and would have had a reasonable expectation of success in using, standard deviation to calculate Takano’s threshold” because doing so “would have been combining prior art elements (difference parameters and a standard deviation) according to known methods (calculating standard deviations) to yield predictable results (standard deviation of the speech signals to generate a gain threshold).” *Id.* ¶ 126.

At this stage of the proceeding, Patent Owner makes no arguments against combining Lyons’s teachings with the teachings of Holzrichter, Takano, and Kinsler. *See, e.g.*, Prelim. Resp. 14. Nonetheless, the burden

remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record, Petitioner establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had reasons, e.g., as articulated by Dr. Stern, to combine Lyons’s teachings with the teachings of Holzrichter, Takano, and Kinsler. *See* Pet. 44–46; Ex. 1002 ¶¶ 122–126. Petitioner also establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had a reasonable expectation of success in combining the teachings of the references. *See* Pet. 46; Ex. 1002 ¶ 126.

3. *Preliminary Conclusion About Obviousness/Nonobviousness*

As discussed above, Petitioner’s analysis explains how the combined disclosures in Holzrichter, Takano, Kinsler, and Lyons teach claim 3’s subject matter. *See supra* §§ III.E.1.a–III.E.1.i. Additionally, Petitioner provides a reason with rational underpinning as to why an ordinarily skilled artisan would have had reason to combine Lyons’s teachings with the teachings of Holzrichter, Takano, and Kinsler and would have had a reasonable expectation of success. *See supra* § III.E.1.j. Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claim 3 unpatentable under § 103(a) as obvious over Holzrichter, Takano, Kinsler, and Lyons.

F. Alleged Obviousness over Holzrichter, Torick, Kaneda, and Kinsler: Claims 1, 2, 4, and 5

Petitioner contends that claims 1, 2, 4, and 5 are unpatentable under § 103(a) as obvious over Holzrichter, Torick, Kaneda, and Kinsler. *See* Pet. 14, 47–55. Patent Owner disputes Petitioner’s contentions. *See* Prelim.

Resp. 14–17. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Torick, Kaneda, and Kinsler teach the subject matter in claims 1, 2, 4, and 5.

1. Analysis of Independent Claim 1

a) Introduction

For the challenge based on Holzrichter, Torick, Kaneda, and Kinsler, Petitioner relies on Holzrichter and Kinsler in the same way as for the challenge based on Holzrichter, Takano, and Kinsler to teach the following in claim 1: the preamble, the “voicing sensor” limitation, the “processor” limitation, the “cross correlation data” limitation, and the “identifies voiced speech” limitation. Pet. 47; *see id.* at 18, 26–35; *supra* §§ III.D.4(a), III.D.4(c)–III.D.4(f). The analysis below discusses Petitioner’s reliance on Holzrichter, Torick, and Kaneda to teach claim 1’s other limitations and the motivation to combine the references. *See infra* §§ III.F.3(b)–III.F.3(f).

b) At Least Two Microphones

Claim 1 recites “at least two microphones that receive the acoustic signals.” Ex. 1001, 11:11.

Petitioner contends that the combined disclosures in Holzrichter and Torick teach claim 1’s “two microphones” limitation. *See* Pet. 47–49.

Specifically, Petitioner asserts that:

- (1) “Holzrichter discloses using the strength of the output signal of a microphone to determine if a signal is speech or noise”; and
- (2) “Torick discloses a system for detecting ‘speech sounds’ using two microphones,” where one microphone detects speech and the other microphone detects ambient noise.

Id. at 47.

Petitioner also asserts that “Torick uses these two microphones to implement ‘voice actuated switching’ and ‘eliminate the press-to-talk switch’ for microphone-based systems.” Pet. 48 (citing Ex. 1002 ¶ 132; Ex. 1006, 168).

At this stage of the proceeding, Patent Owner does not dispute that (1) Holzrichter discloses a one-microphone system and (2) Torick discloses a two-microphone system. *See* Prelim. Resp. 12–18.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter and Torick teach claim 1’s “two microphones” limitation. *See* Pet. 47–49; Ex. 1002 ¶¶ 130–133. Specifically, Torick discloses a two-microphone system with a noise-cancelling speech microphone to detect speech and a noise-sensing microphone to detect ambient noise. Ex. 1006, 168–69; *see* Ex. 1002 ¶ 131. The system uses the two microphones to implement voice-actuated switching (VOX) and “eliminate the press-to-talk switch.” Ex. 1006, 168–69; *see* Ex. 1002 ¶ 132. The system provides “highly reliable voice-actuated switching (VOX) at widely varying ambient levels” and permits operation in high ambient noise. Ex. 1006, 168; *see* Ex. 1002 ¶ 132.

c) Processor Generates Difference Parameters

Claim 1 specifies that the processor “generates difference parameters between the acoustic signals received at each of the two receivers, wherein the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals.” Ex. 1001, 11:25–29.

Petitioner contends that “Holzrichter combined with Torick in light of Kaneda” teaches claim 1’s “difference parameters” limitation. *See* Pet. 51–52. Specifically, Petitioner asserts that Torick discloses the following:

- (1) the “noise-sensing microphone output is amplified and equalized to correspond to the far-field response of the noise-cancelling speech microphone”;
- (2) the system continuously compares the “outputs of both microphones”;
- (3) when the “outputs of both microphones” are equal, “the switching circuit is inactive”; and
- (4) when “speech sounds are present at the appropriate microphone, the switching circuit energizes.”

Id. at 51 (quoting Ex. 1006, 169).

Petitioner asserts that it “would have been obvious to” an ordinarily skilled artisan reading Torick that, “in order for the system to function properly, there should be some tolerance or range within which Torick would consider the outputs of two microphones to be equal and outside of which Torick would consider the outputs to not be equal, thereby detecting speech.” Pet. 51 (citing Ex. 1002 ¶ 142). Petitioner also asserts that an ordinarily skilled artisan would have looked for “methods to compare the outputs of the two signals and determine whether the two signals are within a certain range of each other.” *Id.* at 51–52 (citing Ex. 1002 ¶ 142).

Additionally, Petitioner contends that “Kaneda teaches that one way to compare the outputs of two microphones is to calculate a ‘power difference PD’ (difference parameters) by subtracting the power of one microphone from the other.” Pet. 52 (citing Ex. 1002 ¶ 143; Ex. 1007, 2:1–22, Figs. 3(a)–3(c)). According to Petitioner, “[i]t would have been

obvious to compare the outputs of Torick’s microphones in the manner taught by Kaneda because subtracting one signal from another is a standard technique for comparing two signals.” *Id.* (citing Ex. 1002 ¶ 144).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1’s “difference parameters” limitation. *See, e.g.*, Prelim. Resp. 14–17. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Torick, and Kaneda teach claim 1’s “difference parameters” limitation. *See* Pet. 51–52; Ex. 1002 ¶¶ 140–144.

*d) Processor Identifies Information of the
Acoustic Signals as Unvoiced Speech*

Claim 1 specifies that the processor “identifies information of the acoustic signals as unvoiced speech when the difference parameters exceed a gain threshold.” Ex. 1001, 11:30–32.

Petitioner contends that the combined disclosures in Holzrichter, Torick, and Kaneda teach claim 1’s “identifies unvoiced speech” limitation. *See* Pet. 53–54. Specifically, Petitioner asserts that:

- (1) Holzrichter discloses identifying a signal “as unvoiced speech when no vocal motion is present and the output of signal is sufficient to identify the signal as speech”;
- (2) Torick discloses detecting “speech sounds” when the output of the noise-cancelling speech microphone exceeds the output of the noise-sensing microphone “analogous to the presence of a virtual threshold”; and

- (3) Kaneda discloses “identifying a portion of a signal as speech when ‘[w]hen the power difference PD is larger than a given threshold value Pth17.’”

Id. at 53 (alteration by Petitioner) (emphasis omitted) (quoting Ex. 1006, 169; Ex. 1007, 2:15–17) (citing Ex. 1002 ¶¶ 146–147; Ex. 1003, 31:61–32:27; Ex. 1007, 2:13–18).

Additionally, Petitioner contends that “the combined system would identify a signal as unvoiced speech only when there is no vocal motion present and when the difference in outputs between Torick’s two microphones exceeds a threshold.” Pet. 54 (emphasis omitted) (citing Ex. 1002 ¶ 148).

At this stage of the proceeding, Patent Owner makes no arguments specific to claim 1’s “identifies unvoiced speech” limitation. *See, e.g.*, Prelim. Resp. 14–17. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Torick, and Kaneda teach claim 1’s “identifies unvoiced speech” limitation. *See* Pet. 53–54; Ex. 1002 ¶¶ 145–148.

*e) Processor Identifies Information
of the Acoustic Signals as Noise*

Claim 1 specifies that the processor “identifies information of the acoustic signals as noise when the difference parameters are less than the gain threshold.” Ex. 1001, 11:33–35.

Petitioner contends that the combined disclosures in Torick and Kaneda teach claim 1’s “identifies noise” limitation. *See* Pet. 54–55.

Specifically, Petitioner asserts that Torick discloses not detecting “speech sounds” when the output of the noise-cancelling speech microphone “is the same as” the output of the noise-sensing microphone. *Id.* at 54 (citing Ex. 1002 ¶ 150; Ex. 1006, 169). Petitioner asserts that “Kaneda similarly discloses identify a portion of a signal as noise when the power difference PD is lower than a threshold.” *Id.* (citing Ex. 1002 ¶ 150; Ex. 1007, 1:51–2:22, Figs. 3(a)–3(c)).

Patent Owner contests Petitioner’s position and contends that

As with Ground 1A discussed above, Petitioner’s arguments boil down to conflating the ***absence*** of an identification of speech with an identification of acoustic signals as noise. These arguments fail for largely the same reasons, and the Petition should be denied.

Prelim. Resp. 8–9. Patent Owner specifically argues that Petitioner appears to misquote Torick and that when speech sounds are present in Torick, then its switching circuit is energized. Prelim. Resp. 13 (citing Ex. 1006, 169). Thus, according to Patent Owner, Torick merely states that a “switching circuit energizes” when “speech sounds are present at the appropriate microphone.” *Id.* But, Patent Owner asserts that “the ***absence*** of activity due to speech not being present in a microphone cannot constitute an identification of a signal as noise.” *Id.* Rather, Torick does not disclose identifying *any* signal as anything when its switching circuit is inactive. *Id.* at 13–14.

Patent Owner then argues that “Kaneda discloses a system that detects a ‘speech period’ if a power difference is larger than a threshold—it does **not** disclose detecting any equivalent ‘noise period,’ let alone identifying an acoustic signal as noise.” *Id.* at 14–15.

Lastly, Patent Owner contends that Petitioner mischaracterizes Holzrichter, because it does not disclose identifying any acoustic signal as noise. Prelim. Resp. 15–16. To the contrary, Patent Owner argues that Holzrichter teaches that if “no vocal fold motion occurs in box 72, the sound in frame ti is labeled as **nonvoiced speech** in box 79” (as opposed to noise). On the other hand “[i]f the acoustic signal is too loud in box 77, and vocal fold motion is detected at box 80, the signal is processed as **loud voiced speech**.” *Id.* at 16.

At this stage and based on the current record, we are persuaded Petitioner has shown sufficiently that the combined teachings of Holzrichter, Torick, and Kaneda teach the challenged limitation. As noted regarding the previous challenge, we preliminarily determine that claim 1 sets the boundary for what is considered “noise.” Specifically, claim 1 recites that if the “difference parameters are less than the gain threshold” then the system “identifies information of the acoustic signal as noise.” *See* Ex. 1001, 11:33–35. Therefore, anything that is less than the gain threshold is considered “noise” by the ’058 patent even if its silence.

Accordingly, based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Torick and Kaneda teach claim 1’s “identifies noise” limitation. *See* Pet. 54–55; Ex. 1002 ¶¶ 149–151.

f) Alleged Reasons for Combining the Teachings of the References

Petitioner identifies reasons that would have prompted an ordinarily skilled artisan to combine the teachings of Holzrichter, Torick, and Kaneda in the way Petitioner proposes. *See* Pet. 49–53. Additionally, Petitioner

explains that an ordinarily skilled artisan “would have had a reasonable expectation of success” in combining the teachings of the references. *Id.* at 50. Dr. Stern’s testimony supports Petitioner’s positions. *See* Ex. 1002 ¶¶ 132–140, 144, 148.

For combining the teachings of Holzrichter and Torick, Dr. Stern cites Holzrichter’s disclosure that “the strength of acoustic signals could be used to determine noise or speech.” Ex. 1002 ¶ 135 (citing Ex. 1003, 31:67–32:10). He explains that an ordinarily skilled artisan would have been motivated by that disclosure “to look for other methods for classifying signals as noise or speech based on signal strength.” *Id.* He also explains that “Torick discloses such as a system” with an “improved method of distinguishing speech from noise by comparing the outputs [of] two microphones.” *Id.* ¶¶ 133, 135.

Dr. Stern testifies that adding another microphone to Holzrichter’s single-microphone system according to Torick’s teachings would advantageously:

- (1) improve performance by better distinguishing “between noise and speech”;
- (2) improve performance by allowing “for the accurate identification of speech even in ‘high noise environments’”;
- (3) “more accurately identify an acoustic signal as speech for further processing”; and
- (4) conserve “computing resources” by permitting the system to “perform additional processing only on audio samples that are more likely to be speech.”

Ex. 1002 ¶¶ 136, 138, 140 (citing Ex. 1006, 169); *see id.* ¶ 148. Dr. Stern also testifies that “the proposed combination represents nothing more than a

predictable use of prior art elements according to established functions (i.e., using two microphones to detect speech).” *Id.* ¶ 137.

For combining the teachings of Torick and Kaneda, Dr. Stern testifies that an ordinarily skilled artisan “would have been motivated to combine Kaneda’s method with Torick because it would allow” adding “sufficient tolerance to Torick’s system to allow the system to operate.” Ex. 1002 ¶ 144. He also testifies that “the proposed modification represents nothing more than the simple addition of one known element (Kaneda’s subtraction) to another known element (Torick’s microphones) to obtain” the “predictable result of comparing the two signals from Torick’s microphones.” *Id.*

Based on the current record, we are persuaded that Petitioner establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had reasons to combine the teachings of Holzrichter, Torick, and Kaneda in the way Petitioner proposes. *See* Pet. 49–52; Ex. 1002 ¶¶ 132–140, 144, 148. For combining the teachings of Holzrichter and Torick, Dr. Stern testifies that adding another microphone to Holzrichter’s single-microphone system according to Torick’s teachings would advantageously improve performance by better distinguishing “between noise and speech” and by allowing “for the accurate identification of speech even in ‘high noise environments.’” Ex. 1002 ¶¶ 136, 138, 140; *see id.* ¶¶ 132–133; Pet. 49–50.

Based on the current record, we are persuaded that improving performance suffices as a reason for combining the teachings of Holzrichter and Torick. *See* Ex. 1002 ¶¶ 133, 136, 138, 140. 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” references. *KSR*, 550 U.S. at 420. For combining the teachings of Torick and Kaneda, Dr. Stern testifies that an ordinarily skilled artisan “would have been motivated to combine Kaneda’s method with Torick because it would allow” adding “sufficient tolerance to Torick’s system to allow the system to operate.” Ex. 1002 ¶ 144; *see* Pet. 52. Based on the current record, we are persuaded that adding “sufficient tolerance to Torick’s system to allow the system to operate” suffices as a reason for combining the teachings of Torick and Kaneda. *See* Ex. 1002 ¶ 144.

g) Preliminary Conclusion for Obviousness/Nonobviousness of Claim 1

As discussed above, Petitioner’s analysis explains how the combined disclosures in Holzrichter, Torick, Kaneda, and Kinsler teach claim 1’s subject matter. *See supra* §§ III.F.1(b)–III.F.1(e). Additionally, Petitioner provides a reason with rational underpinning as to why an ordinarily skilled artisan would have been motivated to combine the teachings of Holzrichter, Torick, Kaneda, and Kinsler in the way Petitioner proposes and would have had a reasonable expectation of success. *See supra* § III.F.1(f). Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claim 1 unpatentable under § 103(a) as obvious over Holzrichter, Torick, Kaneda, and Kinsler.

2. Analysis of Independent Claim 2

Independent claim 2 recites “[a] method for removing noise from acoustic signals” comprising steps that parallel the processor operations specified in claim 1. Ex. 1001, 12:1–24.

Petitioner asserts that claim 2 is “similar to the system claimed in claim 1” and obvious over Holzrichter, Torick, Kaneda, and Kinsler “for the same reasons as claim 1.” Pet. 55. Dr. Stern’s testimony supports Petitioner’s assertions. *See* Ex. 1002 ¶ 152.

Patent Owner does not dispute that the references render claim 2 obvious. *See generally* Prelim. Resp. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

For the reasons discussed above, we consider Patent Owner’s arguments regarding claim 1 unavailing. *See, e.g., supra* §§ III.F.1(g). Hence, for the reasons discussed above for claim 1, Petitioner demonstrates a reasonable likelihood of prevailing in proving claim 2 unpatentable under § 103(a) as obvious over Holzrichter, Torick, Kaneda, and Kinsler. *See supra* §§ III.D.1(b)–III.D.1(f).

3. *Analysis of Dependent Claims 4 and 5*

Claims 4 and 5 depend directly from claim 2. Ex. 1001, 12:28–33.

For claims 4 and 5, Petitioner provides an analysis supported by Dr. Stern’s testimony about how Holzrichter teaches the limitations in claims 4 and 5. *See* Pet. 41–43, 55; Ex. 1002 ¶¶ 112–115, 153–154.

At this stage of the proceeding, Patent Owner makes no arguments specific to claims 4 and 5. *See, e.g.,* Prelim. Resp. 14–17. Nonetheless, the burden remains on Petitioner to demonstrate unpatentability. *See Dynamic Drinkware*, 800 F.3d at 1378.

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that Holzrichter teaches the

limitations in claims 4 and 5. *See* Pet. 41–43, 55; Ex. 1002 ¶¶ 112–115, 153–154. Hence, Petitioner demonstrates a reasonable likelihood of prevailing in proving claims 4 and 5 unpatentable under § 103(a) as obvious over Holzrichter, Torick, Kaneda, and Kinsler.

G. Alleged Obviousness over Holzrichter, Torick, Kaneda, Kinsler, and Lyons: Claim 3

Petitioner contends that claim 3 is unpatentable under § 103(a) as obvious over Holzrichter, Torick, Kaneda, Kinsler, and Lyons. *See* Pet. 14, 55–56. Patent Owner disputes Petitioner’s contentions. *See* Prelim. Resp. 18. Above, we provided overviews of Holzrichter, Torick, Kaneda, Kinsler, and Lyons. *See supra* §§ III.F.1. As explained below, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Torick, Kaneda, Kinsler, and Lyons teach claim 3’s subject matter.

Claim 3 depends directly from claim 2 and further requires “generating the gain threshold using standard deviations corresponding to the generation of the difference parameters.” Ex. 1001, 12:25–27.

Petitioner contends that the combined disclosures in Torick, Kaneda, and Lyons teach the limitation in claim 3. *See* Pet. 55–56. Specifically, Petitioner asserts that Lyons discloses that “the standard deviation of a signal ‘is an indication of the magnitude of the fluctuating component of the signal.’” *Id.* (quoting Ex. 1008, 479). Petitioner also asserts that it would have been obvious to an ordinarily skilled artisan in view of Lyons to:

- (1) “measure the power differences received by Torick’s microphones in various circumstances when Torick is receiving noise”;

- (2) “calculate the standard deviation of the different power differences, which would represent an upper bound for how high the power difference could reach when receiving noise”; and
- (3) “use this upper bound as the threshold” for determining that the system receives “noise instead of speech.”

Id. (citing Ex. 1002 ¶ 157).

Additionally, Petitioner contends that an ordinarily skilled artisan “would have been motivated to do this” for the reasons discussed for the ground based on Holzrichter, Takano, Kinsler, and Lyons. Pet. 56 (citing Ex. 1002 ¶ 157); *see id.* at 43–47; *supra* § III.F.1.

Patent Owner disputes that Holzrichter, Torick, Kaneda, Kinsler, and Lyons teach claim 3’s subject matter. Prelim. Resp. 18. Specifically, Patent Owner asserts that (1) there are deficiencies in the references as applied to claim 2, and (2) Petitioner’s arguments for claim 3 “do not remedy the deficiencies.” *Id.*

Based on the current record and for the reasons advanced by Petitioner and supported by Dr. Stern’s testimony, Petitioner establishes sufficiently for purposes of institution that the combined disclosures in Holzrichter, Torick, Kaneda, Kinsler, and Lyons teach claim 3’s subject matter. *See* Pet. 56; Ex. 1002 ¶¶ 156–157. Petitioner also establishes sufficiently for purposes of institution that an ordinarily skilled artisan would have had reasons, e.g., as articulated by Dr. Stern, to combine Lyons’s teachings with the teachings of Holzrichter, Torick, Kaneda, and Kinsler. *See* Pet. 44–46, 56; Ex. 1002 ¶¶ 122–126, 157. Petitioner also establishes sufficiently for purposes of institution that an ordinarily skilled

artisan would have had a reasonable expectation of success in combining the teachings of the references. *See* Pet. 46; Ex. 1002 ¶ 126.

Based on the current record and for the reasons discussed above for claims 1 and 2, we disagree with Patent Owner that there are deficiencies in the references as applied to claim 2. *See* Prelim. Resp. 18; *supra* § III.F.1.

IV. CONCLUSION

Based on the arguments and evidence presented by the parties, we determine that there is a reasonable likelihood Petitioner would prevail with respect to at least one claim challenged in the Petition. Hence, we institute an *inter partes* review of all challenged claims on all challenges included in the Petition. *See SAS Inst. Inc. v. Iancu*, 138 S. Ct. 1348, 1355 (2018) (noting that the language of 35 U.S.C. § 314(b) “indicates a binary choice—either institute review or don’t”); *see also PGS Geophysical AS v. Iancu*, 891 F.3d 1354, 1360 (Fed. Cir. 2018) (interpreting the statute as requiring “a simple yes-or-no institution choice respecting a petition, embracing all challenges included in the petition”).

V. ORDER

Accordingly, it is

ORDERED that, pursuant to 35 U.S.C. § 314(a), an *inter partes* review of claims 1–5 in the ’058 patent is instituted on all challenges included in the Petition;

FURTHER ORDERED that, according to 35 U.S.C. § 314(c) and 37 C.F.R. § 42.4, notice is hereby given of the institution of a trial that commences on the entry date of this Decision; and

FURTHER ORDERED that the Motion for Joinder is denied.

IPR2023-01110
Patent 7,246,058 B2

For PETITIONER:

Indranil Mukerji
John Moulder
WILLKIE FARR & GALLAGHER LLP
imukerji@willkie.com
cmoulder@willkie.com

For PATENT OWNER:

Peter Lambrianakos
Vincent Rubino
Enrique Iturralde
Richard Cowell
Jacob Ostling
FABRICANT LLP
plambrianakos@fabricantllp.com
vrubino@fabricantllp.com
eiturralde@fabricantllp.com
rcowell@fabricantllp.com
jostling@brownrudnick.com