

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM, INC.,
AMAZON.COM SERVICES LLC, and
AMAZON WEB SERVICES, INC.,
Petitioners,

v.

SOUNDCLEAR TECHNOLOGIES LLC,
Patent Owner.

Case No. IPR2025-00565
U.S. Patent No. 11,069,337

PATENT OWNER'S SUR-REPLY

TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION	1
II. CLAIM CONSTRUCTION	3
A. No New Construction Is Permissible	3
B. Distance Does Not Equal Voice.....	3
III. GROUNDS 1-5 FAIL TO PROVE BY A PREPONDERANCE OF THE EVIDENCE THAT ANY CHALLENGED CLAIMS ARE UNPATENTABLE.....	5
A. None of the Asserted References Disclose the Claimed “Voice” Classification (Grounds 1-5).....	5
B. None of the References Disclose the Claimed “Classification” (Grounds 1-5)	5
1. Unsupported Assumptions Based on Adjusted Outputs Do Not Suffice to Demonstrate Disclosure of the Claimed “Classification” Limitation	5
2. At Best, the Asserted References Classify <i>Distance</i>	10
C. No Analysis of Voice	11
1. Shin Does Not Disclose “Analyze a Voice ... to Classify the Voice” (Grounds 1 and 2).....	11
2. Shimomura Does Not Disclose “Analyze a Voice ... to Classify the Voice” and Related Limitations (Ground 3).....	16
D. Obviousness Grounds Fail	18
1. Articulated Reasoning with Rational Underpinnings Required	18
2. Adjusting of the Sound Volume of Voice Data Was Not Obvious (Grounds 3-5)	19

3.	Shimomura in View of Shin (Ground 4) Fails.....	21
IV.	CONCLUSION.....	22

TABLE OF AUTHORITIES

	Page(s)
Cases	
<i>Axonics, Inc. v. Medtronic, Inc.</i> , 75 F.4th 1374 (Fed. Cir. 2023)	3
<i>Crown Operations Int’l, Ltd. v. Solutia Inc.</i> , 289 F.3d 1367 (Fed. Cir. 2002)	6, 8, 9
<i>Graham v. John Deere Co.</i> , 383 U.S. 1 (1966).....	18
<i>Johns Manville v. Knauf Insulation, inc. et. al.</i> , IPR2018-00827, Paper 9, 11-12 (PTAB Oct. 16, 2018).....	18
<i>KSR Int’ l Co. v. Teleflex Inc.</i> , 550 U.S. 398, 127 S. Ct. 1727 (2007).....	18, 19, 20
<i>Net MoneyIN, Inc. v. VeriSign, Inc.</i> , 545 F.3d 1359 (Fed. Cir. 2008)	15
<i>Phillips v. AWH Corp.</i> , 415 F.3d 1303 (Fed. Cir. 2005)	3
<i>Rembrandt Diagnostics, LP v. Alere, Inc.</i> , 76 F.4th 1376 (Fed. Cir. 2023)	14
<i>Yita LLC v. MacNeil IP LLC</i> , 69 F.4th 1356 (Fed. Cir. 2023)	14
Statutes	
5 U.S.C. § 554(b)-(c)	3
Other Authorities	
37 C.F.R. § 42.23	14

PATENT OWNER'S EXHIBIT LIST

Exhibit No.	Description
2001	Docket Printout of CA No. 1:24-01283 (EDVA)
2002	Order from Judge Anthony Trenga, dated January 10, 2025 (Dkt. 84)
2003	Scheduling Order, dated September 23, 2024 (Dkt. 34)
2004	Final Scheduling Order, dated October 18, 2024(Dkt. 55)
2005	Judge Anthony Trenga's median time to trial statistics
2006	Excerpts from the Motion to Stay hearing transcript, dated March 26, 2025
2007	Joint proposed scheduling order (Dkt. 48)
2008	Reply in Support of Motion to Stay, dated March 20, 2025 (Dkt. 129)
2009	Full prosecution history of U.S. Patent No. 11,069,337
2010	TransPerfect Blog – The Evolution of AI in Language Processing and Its Impact on IP/Patents
2011	Median Time to Disposition in Cases Terminated After Hearing or Submission, UNITED STATES COURT OF APPEALS FOR THE FEDERAL CIRCUIT, available at https://www.cafc.uscourts.gov/wp-content/uploads/reportsstats/FY2024/MedianDispositionTime-Merits.pdf .
2012	Declaration of Anderson To
2013	Declaration of Patent Owner Expert Joseph C. McAlexander
2014	Curriculum Vitae of Joseph C. McAlexander
2015	Transcript of October 30, 2025 Deposition of Petitioners' Expert, Dr. Richard Stern
2016	Transcript of April 8, 2026 Deposition of Petitioners' Expert Dr. Richard Stern

I. INTRODUCTION

The Petition fails for three independent reasons. First, none of the prior art discloses classification of “voices.” Second, Petitioners’ expert admitted that there is no basis on which to conclude that the prior art discloses “classifying” inputs at all (be it voice, volume, or distance). And third, Petitioners acknowledge—by disclosing a new theory on Reply—that none of the prior art “analyzes a voice spoken by a user ... *to*” do so.

First, none of the prior art discloses classification of “voices.” Petitioners admit that the Petition fails even under the Board’s preliminary construction. That construction would write the concept of “voice” out of the claims. It is impermissibly broad. And yet it is still not broad enough to save the Petition. That is why Petitioners contest it. They need the Board to go even further, to construe different “voices” to mean different “distances.” Doing so would create an implausible result. That interpretation would read on a device that adjusts output whenever distance changes—even if the device never attempts to classify a voice.

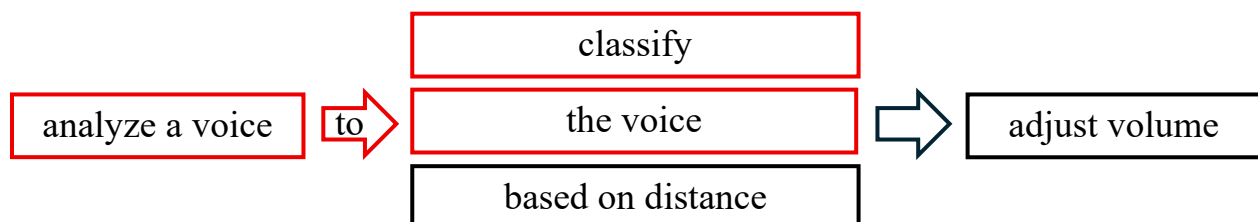
Second, Petitioners’ expert admitted that he baselessly assumed that any input variable that causes a change in output has been classified. He also testified that doing so is indefensible. Indeed, attempting to read the prior art onto the claims in such a way creates implausible results. Petitioners could use that backward assumption to argue that the prior art classifies almost any input variable

to the device, such as “words,” “frequency,” “volume,” “emotion,” or “user intent.”

Or they could go even further. They could argue that the prior art classifies other things affecting distance, such as “identity of the user” (e.g., which changes when a person in a different location speaks), “location of the user,” “motion of the user” (e.g., backward or forward), or even whether the user is doing a summersault. That is because changing any of those latter variables would also change the distance measurement. And each such change would have the *exact same* effect on altering the outputs of the cited prior art. It is improper to interpret and apply prior art in a manner that contradicts its disclosure and yields a result that is implausible.

Third, none of the prior art “analyzes a voice spoken by a user ... *to*.” The claims recite analysis of a voice for a purpose. That purpose is “to” classify it. Simply measuring distance with a proximity sensor is entirely different. It is a measurement, not a classification. And the object is distance, not voice.

In short, the Petition cannot succeed unless the following aspects of the independent claims shown in red are ignored or rewritten.



II. CLAIM CONSTRUCTION

A. No New Construction Is Permissible

Both parties agree that the preliminary construction is wrong. Reply, 5; POR, 5-21. Pivoting to a new construction now would violate the Administrative Procedure Act, 5 U.S.C. § 554(b)-(c); *Axonics, Inc. v. Medtronic, Inc.*, 75 F.4th 1374, 1383 (Fed. Cir. 2023). It would also be futile, since there is no argument that any prior art discloses classification of a “voice.”

B. Distance Does Not Equal Voice

The claims require classification of a “voice.” EX-1001, cls. 1[b], 4[c], 5[c]. The patentee chose the clearest, most straightforward language available to convey that the voice is the object that must be classified: “classify the voice.”¹

The parties agree that “voice” is distinct from “volume” and “distance.” Distance (e.g., 5 feet) is not a voice (e.g., a whisper). But regardless, both experts’ testimony is unrebutted on this front.

Q Does the distance between the user and the device alone tell you anything about the characteristics of the user’s voice?

¹ It is “unjust to the public, as well as an evasion of the law, to construe [an invention] in a manner different from the plain import of its terms.” *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312 (Fed. Cir. 2005).

MR. REEVES: Objection. Form. Outside the scope.

A In general, no.

EX-2015, 54:1-6; *see also id.*, 28:12-18 (“The voice I consider to be the natural productions that humans make in communicating with each other via speech.”); 31:11-32:6 (“I’m talking about the subset of things that humans utter that are directed toward communication of an idea or a request or something like that, normal conversation.”); EX-2013, ¶¶44, 53-55 *see also generally id.*, ¶¶44-55; Dec., 16 (users “perform” voices).

Petitioners tacitly admit that every ground in the Petition fails without a construction. And not only that, they admit the Petition fails unless the Board goes even further. Petitioners object to the preliminary construction: “To the extent the Board found the claims to require change in distance and volume, Petitioner agrees with Patent Owner that *this would be incorrect.*” Reply, 5 (emphasis altered). They argue that “voice” must be read to exclude any notion of “volume.” *Id.* That is because none of the prior art classifies volume—let alone a voice.

Petitioners cannot escape that the preliminary construction requires “a single user who may perform both a first and second voice depending upon the single user’s changing distance (*and associated changing voice volume*) from the voice-content control device.” Dec., 16 (emphasis added). The Board determined that any changing volume could constitute a changing voice and that detecting that change

in distance constitutes a change in voice. The Board made that clear when it included the word “volume” in the construction and tied it to distance using the word “associated.”

Even if it were proper to revisit claim construction, Petitioners could not succeed. The plain meaning of “classify the voice” cannot be expanded to “classify the distance.” And it would lead to implausible results. *See infra* Section III.B.1.

III. GROUNDS 1-5 FAIL TO PROVE BY A PREPONDERANCE OF THE EVIDENCE THAT ANY CHALLENGED CLAIMS ARE UNPATENTABLE

A. None of the Asserted References Disclose the Claimed “Voice” Classification (Grounds 1-5)

None of the prior art discloses classifying a “voice.” Both experts testified that the concept of a “voice” is separate from the concepts of “volume” or “distance.” *See supra* Section II.B. Any construction or meaning applied by the Board that only requires distance classification would be wrong under the law and would conflict with the factual record (including the testimony of both experts).

B. None of the References Disclose the Claimed “Classification” (Grounds 1-5)

1. Unsupported Assumptions Based on Adjusted Outputs Do Not Suffice to Demonstrate Disclosure of the Claimed “Classification” Limitation

The Petition does not identify any disclosure in the prior art of the recited voice classification. It relies solely on unsupported assumptions based on inputs

and outputs. Pet., 12-14, 39-40. That is improper. *Crown Operations Int'l, Ltd. v. Solutia Inc.*, 289 F.3d 1367, 1377 (Fed. Cir. 2002) (assumptions about operation of prior art improper). As much is clear in the concluding sentences of the applicable paragraphs. Those are the instances in which the Petition directly addresses voice “classif[ication].” Pet., 13 (first full paragraph, last sentence); *id.*, 39 (same). Petitioners do so in purely conclusory fashion. *Id.* Neither portion cites any disclosure in the prior art references. *Id.* They both rely solely on the same conclusory expert statement. And that statement mirrors the attorney argument in the Petition in both instances. *Compare id.*, 13 (“To do so, Shin’s processor classifies the voice based on the distance.”) (citing EX-1002 ¶63)), 39 (“To do so, Shimomura’s control unit first classifies the voice.” (citing EX-1002 ¶138)) *with* EX-1002, ¶¶63, 138. The record is bereft of any analysis establishing that Shin or Shimomura actually “do so” when it comes to classification.

Dr. Stern’s testimony confirms that adjusting an output based on an input does *not* meet the claims:

Q I’m going to -- I’ll call it a new motivational output-content control device, and it’s pretty simple. When a user asks a question, it measures the volume detected, and then it uses, you know, a language model to develop an answer, and it responds at twice the volume at which the user was speaking.

So my logic would say output volume of the device, call it OV, equals two times the input volume of the device, IV.

In that hypothetical, is my device classifying input voices?

MR. REEVES: Objection. Form. Calls for speculation. Incomplete hypothetical. Outside the scope.

THE WITNESS: No.

EX-2016, 64:5-65:2.

In the excerpt above, Dr. Stern was asked about a hypothetical device that measures an input volume and doubles it to create an output volume. He agreed that that would *not* be a classification of anything. *Id.* He testified at most about classifying outputs as a separate step: “classifying anything as I see it, describes taking an input measurement that, in general, is continuous in nature and producing an output that’s one of a discrete number of categories in the patents in question.” *Id.* 65:7-11 (emphasis added). Even then, he confirmed that the input—the speaker’s distance—remains continuous: “Is the speaker closer or father.” *Id.* 67:11-12 (emphasis added). Thus, we return to Dr. Stern’s opinion that measuring a continuous value is not a classification. Classifying output—the results—does not mean that the inputs were classified.

Here again, Dr. Stern agrees. Analyzing outputs tells one nothing about what a device does in terms of processing inputs:

Q ... Would a POSITA have been able to determine how a device classifies inputs simply by looking at how the device adjusts outputs?

MR. REEVES: Objection. Form. Incomplete hypothetical. Outside the scope.

THE WITNESS: I think, in general, the answer would be no. I don't believe that that's relevant to the claims of the present patent or the '337 patent.

EX-2016, 62:22-63:8.

He also testified that making assumptions based upon inputs and outputs is still insufficient and indefensible when assessing whether a classification occurs. He stated that one must examine the inner workings of a device to identify a classification—not merely infer it from looking at inputs and outputs. *Id.*, 33:5-34:15 (even knowing the inputs, one would have to examine the “code or the hardware implementation” to determine what a classification is based on); 77:19-78:3 (analyzing operations of algorithms for something as simple as thresholds would require looking at code or hardware); *id.*, 21:8-22:16 (those are “black box” functionalities). For the same reason, a POSITA would not assume that a classification necessarily occurs. An express or inherent disclosure is required in the prior art. *Crown Operations*, 289 F.3d at 1377.

That makes sense. One cannot assume that any input variable that causes a change in output has been classified. For example, attempting to read the prior art onto the claims in such a way creates implausible results. Petitioners could use that backward assumption to argue that the prior art classifies almost any input variable to the device, such as “words,” “frequency,” “volume,” “emotion,” or “user intent.” Or they could go even further. They could argue that the prior art classifies other things affecting distance, such as “identity of the user” (e.g., which changes when a person in a different location speaks), “location of the user,” “motion of the user” (e.g., backward or forward), or even whether the user is doing a summersault. That is because changing any of those latter variables would also change the distance measurement. And each such change would have the *exact same* effect on altering the outputs of the cited prior art.

It is improper to interpret and apply prior art in a manner that contradicts its disclosure and yields a result that is illogical or implausible. Thus, even *if* the preliminary construction were accepted, and changing voice were the same as changing volume or distance, Dr. Stern's testimony *still* confirms that the Petition insufficiently addresses the recited classification. *Id.*

In short, Petitioners identify no express disclosure of any classification of inputs in Shin or Shimomura. The claim limitation is completely missing from the asserted references. Accordingly, the Petition has not satisfied its burden to show

that the asserted references disclose “classify[ing] the voice as either one of a first voice or a second voice” as required by the claims. EX-1001, cls. 1[b], 4[c], 5[c].

2. At Best, the Asserted References Classify *Distance*

The asserted references, if anything, classify distances, not voices as required by the challenged claims. *See* EX-1001, cl. 1 (“a voice classifying unit configured to analyze a voice spoken by a user and acquired by a voice acquiring unit to classify *the voice*.”) (emphasis added).

Petitioners agree that the Petition relies on classifications of distance. The Reply identifies distance values as the product of the classification: “The left column [of Shin Table 3] includes *categories* (or classifications) for the distance D.” Reply, 11 (emphasis added). The Reply makes the same argument with respect to Shimomura. Reply, 18 (identifying two distance values in Figure 7).

Further, Petitioners’ newly proposed construction removes all doubt that the Petition relies on classification of outputs. It proposes that “classify” be construed to mean “arrange in categories.” Reply, 4. But, as discussed above (Reply, 11), the “categories” Petitioners identify are categories of distance. Distance is what is arranged, what is categorized.

Petitioners’ expert, Dr. Stern, repeatedly states that the claims require voice classification. EX-1202, ¶¶14-16. But all testimony that the prior art classifies *voices* (e.g., *id.*, ¶¶19-20, 24-25) is conclusory and directly contradicts more

specific testimony. IPR2025-00673, EX-2017, 48:3-7 (Q: “Table 3 classifies the *distances* between a user and the electronic device, correct?” A: “Yes.”), 49:7-15 (Q: “Table 3 does not mention voices at all, correct?” A: “The table itself indeed only talks about the distance. ... [I]t is correct that Table 3 does not refer to voices explicitly”); *see also id.*, 117:1-21, 118:2-13, 119:6-22.

Thus, Petitioners admit that the only *possible* classification disclosed in the prior art is of distance. Because the challenged claims require classification of voices and the asserted references at best disclose classification of distances, Petitioners cannot meet their burden to show any challenged claims are unpatentable by a preponderance of the evidence.

C. No Analysis of Voice

1. Shin Does Not Disclose “Analyze a Voice ... to Classify the Voice” (Grounds 1 and 2)

Shin discloses only a limited form of acoustic analysis at the input stage. Shin calls it a “voice analysis,” but Petitioners avoid analyzing how it actually functions and what purpose it serves. The only purpose is to generate output content. POR, 23-24. It has nothing to do with Petitioners’ alleged classification (of distance)—much less the voice classification claimed in the ’337 patent. POR, 23-24, 29-30.

Petitioners admit that Shin discloses “receiv[ing] a speech of a user as a voice input ... *to determine an output scheme*” Reply, 9 (quoting EX-1003, Abstract) (emphasis changed); *see also id.* (“Shin discloses analyzing a voice input ... to alter output.”). And none of the disclosures in Shin cited by Petitioners (nor anywhere else in Shin) discloses analyzing a voice input to *classify the voice*. *Id.* (citing EX-1003, Abstract, ¶¶[0003]-[0011], [0037]-[0039], [0041], [0060]-[0061], [0155], [0158], cls. 1, 15, 20); *see also id.*, 14-15 (citing Pet., 12-13); Pet., 12-13 (citing EX-1003, Abstract, ¶¶[0003]-[0011], [0037]-[0039], [0041], [0047], [0051], [0060]-[0061], [0062], [0077]-[0081], [0086]-[0087], cls. 1, 6, 15, 17, 20).

Petitioners’ Reply does not add any more analysis or argument to support the Petition’s theories. They merely cite back to the Petition and corresponding expert testimony, neither of which include any assertion that Shin analyzes the acquired voice for the purpose of classifying the voice. Reply, 14 (citing Pet., 12-13).

Petitioners’ expert agrees. He repeatedly testified that the primary references (Shin and Shimomura) at best only classify *distances*—not voices. IPR2025-00673, EX-2017, 48:3-7 (Q: “Table 3 classifies the *distances* between a user and the electronic device, correct?” A: “Yes.”) (emphasis added), 49:7-15 (Q: “Table 3 does not mention voices at all, correct?” A: “The table itself indeed only talks about the distance. ... [I]t is correct that Table 3 does not refer to voices explicitly”); *see also id.*, 117:1-21, 118:2-13, 119:6-22. The ’337 patent

specification defines voices as incorporating many characteristics, including volume, frequency, intensity, slopes of peak frequency, content, speaking speed, and others. POR, 15; EX-1001, 7:64-8:34. By way of example, the specification describes classifying “the voice V1 to the second voice V1B when the voice V1 *is determined to be a whisper* as a result of analysis,” and “to the first voice V1A when the voice V1 is *determined not to be a whisper*.” EX-1001, 7:51-63(emphasis added); POR, 15-16 (emphasis changed). This is done by using combinations of “feature values” indicative of the characteristics of the voice. While the specification exemplifies “distance” as one such “feature values[]” of many, the plain language of the claims requires distance to be used in conjunction with other “features values” to classify voices.

Further, Petitioners admit that the claimed “distance between a user and the voice-content control device” is calculated by the “proximity sensor.” *See, e.g.*, EX-1001, cl. 1; Reply, 14-15. Calculating the claimed user-to-device distance with the proximity sensor (as required by limitations 1[a], 4[a], and 5[b]) is not a voice analysis (as required by limitations 1[b], 4[c], and 5[c]). POR, 17-21.

The Petition simply lacks any analysis as to how Shin classifies voices based on voice analysis. Pet., 13-14. This is a foundational problem. POR, 31 n.1.

Petitioners have now acknowledged that problem by attempting to switch to a new (second) theory: “*Second, even if the claims were to require determining distance*

based on voice analysis” Reply, 14-15 (emphasis added). Petitioners now assert that Shin discloses calculating user-to-device distance “by analyzing the voice input.” *Id.* It is too late. That second theory is not mentioned in the Petition. Certainly the Petition does not explain the new mapping or include the newly articulated theory set forth on pages 14 and 15 of the Reply. The theory should thus be rejected. 37 C.F.R. § 42.23; *Rembrandt Diagnostics, LP v. Alere, Inc.*, 76 F.4th 1376, 1383 (Fed. Cir. 2023) (making “a meaningfully distinct contention” or relying on other embodiments for the first time are impermissible); *Yita LLC v. MacNeil IP LLC*, 69 F.4th 1356, 1367 (Fed. Cir. 2023) (affirming the Board’s rejection of a “meaningfully distinct contention” presented for the first time in Petitioner’s Reply).

But even if considered, the Reply’s new theory still fails because the claimed distance calculation and voice analysis are separate limitations. *See* EX-1001, cls. 1, 4-5. There are good reasons to combine user-to-device distance and voice analysis for the purpose of classifying voices. As Petitioners acknowledge, relying on user-to-device distance alone to classify voices will result in different voices being misclassified as the same (e.g., when two users with different voices are at the same distance from the device). Reply, 12. Combining user-to-device distance with voice analysis, as described and claimed by the ’337 patent, mitigates this error. EX-1001, cls. 1[b], 4[c], and 5[c]. And yet Petitioners’ new theory fails to

address the two limitations separately. It relies on the same operation (i.e., calculating user-to-device distance) in Shin to disclose both limitations. It does so without articulating whether or how such double mapping suffices to disclose the limitations of the claims. *Net MoneyIN, Inc. v. VeriSign, Inc.*, 545 F.3d 1359, 1370 (Fed. Cir. 2008) (“[T]he ‘arranged as in the claim’ requirement applies to all claims”).

Finally, Petitioners’ attempt to downplay its own expert’s admissions only highlights the internal contradictions in his testimony. Reply, 12-13. Petitioners’ expert stated that determining whether a classification constitutes a *voice* classification depends on the input and output of the classification. EX-2015, 23:5-24:17. When asked whether Table 3 of Shin “provide[s] any information regarding the user’s voice,” he answered unequivocally: “No.” *Id.*, 72:12-73:8. He further stated that Table 3 of Shin only “tells you about the distance between the user and the electronic device and nothing else.” *Id.* Thus, by Petitioners’ expert’s own logic, Table 3 of Shin does not disclose classification of the user’s voice, because neither the input (i.e., distance) nor the output (device content) of Shin’s Table 3 is the user’s voice. Further, as stated above, Dr. Stern testified that Shin only classifies *distances* and not voices. IPR2025-00673, EX-2017, 48:3-7, 49:7-15.

For these reasons, Shin does not disclose the claimed “analyze a voice ... to classify the voice.”

2. Shimomura Does Not Disclose “Analyze a Voice ... to Classify the Voice” and Related Limitations (Ground 3)

Petitioners argue that “Shimomura analyzes voice (e.g., speech) input *and* classifies it based on distance.” Reply, 18 (citations omitted) (emphasis added). Petitioners’ wording is telling. The Petition’s alleged voice analysis for Shimomura (i.e., speech and sound recognition performed by speech recognition unit 42) has nothing to do with the alleged voice classification (i.e., user-to-device distance determined by image recognition unit 41). POR, 42-44. That is, Shimomura’s analysis of the user’s voice has nothing to do with determination of user-to-device distance (i.e., the basis of the alleged voice classification). Shimomura does not disclose any voice analysis for the purpose of voice classification, as required by the claims.

The Reply does not attempt to show that Shimomura’s alleged voice classification (the alleged determination of user-to-device distance) involves any voice analysis. Reply, 19-20. Nor could they—Petitioners do not and cannot dispute that Shimomura only discloses detecting distance using a camera. *Id.* Petitioners’ sole argument on Reply is that “[n]othing in the claims requires distance to be determined based on a voice analysis.” *Id.* However, the claims require “analyz[ing] a voice ... *to* classify the voice ...based on the distance.” EX-1001, cls. 1[b], 4[c], 5[c] (emphasis added). The detected distance in Shimomura is

different. It has nothing to do with “analyzing the voice ... to classify the voice.”

Shimomura's Figure 7 is only used to change the speech output of Shimomura's robot. EX-1004, ¶¶[0057]-[0058]. Thus, even if Shimomura's classes of distances somehow constituted classes of voices, the Shimomura Grounds still fail because Shimomura's distance classification does not involve any analysis of the acquired voice.

The Shimomura grounds also fail because a POSITA would not have understood Shimomura's classes of distances (described in connection with Figure 7) as disclosing the claimed voice classification. POR, 38-42. On Reply, Petitioners mischaracterize PO's arguments as being based solely in semantics. Reply, 18. They are not merely semantics. Shimomura does not expressly disclose voice classification regardless of word choice. Petitioners' expert agrees. IPR2025-00673, EX-2017, 117:18-21 (Q: “Shimomura illustrates the classification of a distance between a user and the electronic device, correct?” A: That's correct.”), 119:6-22 (Q: “In the context of Shimomura, is there a difference between classifying a voice versus classifying a distance?” A: “Well it's hard for me to answer that, because we never talk about classifying a voice. We talk about classifying based on distance.”).

For these reasons, Shimomura does not disclose the claimed “analyze a voice ... to classify the voice.” For the same reasons, Shimomura likewise does not

disclose the claimed “generate a first [and second] output sentence” based on voice classification.

D. Obviousness Grounds Fail

1. Articulated Reasoning with Rational Underpinnings Required

“[M]ost, if not all” inventions would be deemed obvious if all that were required were disclosure of every claimed element. *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418-419. Assessing obviousness is not an easy task. It *cannot* be dispensed within a sentence or two of MPEP § 2143 incantations.

If obviousness arguments fail to account for hindsight bias, they are per se improper. *KSR*, 550 U.S. at 421 (2007); *Graham v. John Deere Co.*, 383 U.S. 1, 36 (1966) (warning against a “temptation to read into the prior art the teachings of the invention in issue” and instructing courts to “guard against slipping into use of hindsight” (quoting *Monroe Auto Equip. Co. v. Heckethorn Mfg. & Sup. Co.*, 332 F.2d 406, 412 (1964))).

A finding of obviousness requires explicit analysis, including some articulated reasoning with rational underpinnings. *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398 (2007); *Johns Manville v. Knauf Insulation, inc. et. al.*, IPR2018-00827, Paper 9, 11-12 (PTAB Oct. 16, 2018) (informative) (finding no obviousness

where “Petitioner does not support, in its claim-by-claim analysis, which elements of the cited references could be predictably combined.”)

The technology at issue in this IPR is more varied, complex, and unpredictable than the mechanical components in *KSR*. *Id.* at 422-26 (mounting a sensor on a fixed location). And even in *KSR*, the Court discussed the specific facts and obviousness factors at length.

The POR explains the many failings of the Petition. POR, 35-36, 56-62. In short, Petitioners' obviousness grounds are based on conclusory, tautological assertions: that (1) various prior art references disclose the limitations, and (2) combining those references would result in the asserted combination with no analysis to account for hindsight bias. Such statements provide no rational or explanatory basis that accounts for, or prevents, the possibility of hindsight bias.

2. Adjusting of the Sound Volume of Voice Data Was Not Obvious (Grounds 3-5)

Shimomura discloses transforming output content and changing output volume at different distances. POR, 44-46. Specifically, Shimomura discloses, in connection with Figure 7, that the output content should be transformed when user-to-device distance is “350 cm or more.” EX-1004, Fig. 7. In contrast, Shimomura discloses increasing output volume when user-to-device distance is up to “more than 250 [cm]”:

[FIG. 7]

Second Voice	First Voice
Less than 350 cm	350 cm or more
None	Implementation

[FIG. 9]

Less than 50 cm	50~80cm	80~150cm	150~250cm	More than 250
0	+1	+2	+3	+5

52

FIG. 9 Volume Change Table

Compare EX-1004, Fig. 7 with *id.*, Fig. 9; POR, 45-46.

The change in output discussed in Shimomura Fig. 7 is not related to the change in volume discussed in Shimomura Fig. 9. *Id.* Shimomura thus does not expressly disclose adjusting sound volume based on Petitioners' alleged classification (i.e., user-to-device distance more or less than 350 cm).

Petitioners do not dispute that this limitation is completely missing from Shimomura. Instead, they argue that "a POSITA would have been motivated to adjust Shimomura's volume change table, Figure 9, to extend the ranges so they overlapped with Shimomura's output change table, Figure 7." Reply, 20. However, a POSITA would not have been motivated to make this hindsight modification. POR, 44-51. This is because Shimomura expressly teaches transforming output content when further increase in volume is not possible. POR, 47 (quoting EX-1004, ¶¶[0072]-[0073]). Thus, a POSITA would have understood from Shimomura that it is unnecessary to transform output content when increase in output volume

is sufficient for the robot's speech to be understandable by the user. A POSITA would not have been motivated to set the distance at which to change speech volume to be the same as the distance at which to transform the output content. POR, 48-50.

Petitioners fail to respond to PO's argument on Reply. They instead deny that Shimomura teaches simply because it does not "explicitly discourage[] Petitioners' modification." Reply, 21. The mere fact that Shimomura does not explicitly disparage setting the same distance for changing output volume and transforming output content does not change the fact that this proposed modification would directly undermine the embodiments it expressly discloses. Petitioners also assert that Patent Owner's argument rests on "unsupported assumptions." That is not true. Patent Owner's argument is supported by Shimomura's expressed disclosures. POR, 47-51 (quoting EX-1004, ¶¶[0072]-[0073]; citing *id.*, ¶¶[0060]-[0066], [0074]-[0079], [0084]-[0088], Figs. 7, 9-12).

3. Shimomura in View of Shin (Ground 4) Fails

Neither Shimomura nor Shin discloses the claimed volume adjustment. POR, 52-54. Nevertheless, Petitioners argue that this limitation would have been obvious in view of Shimomura and Shin. But as discussed above, a POSITA would not have been motivated to modify Shimomura to set the distance at which to change speech volume to be the same as the distance at which to transform the

output content. *Supra*, Section III.D.2. The same is true even considering Shin, which at best suggests that the user-to-device distance at which to change the output volume and output content *could* be the same as one of an almost unlimited number of options. Regardless, Shimomura still teaches that it is unnecessary to transform output content when increase in output volume is sufficient for the robot's speech to be understandable by the user. Thus, a POSITA would not have been motivated to modify Shimomura to set the same distance for changing output volume and output content.

IV. CONCLUSION

For the foregoing reasons, and those discussed in the POR, the challenged Claims 1-5 of the '337 patent are not unpatentable.

Dated: April 21, 2026

Respectfully submitted,

By: /Steven J. Reynolds/

Steven J. Reynolds (USPTO Reg. No. 61,445)*

Kevin H. Sprenger (USPTO Reg. No. 79,466)

Donald L. Jackson (USPTO Reg. No. 41,090)

Stephen McBride (USPTO Reg. No. 78,396)

Chandran B. Iyer (USPTO Reg. No. 48,434)

Erin Hadi (USPTO Reg. No. 79,904)*

Louay Meroueh (USPTO Reg. No. 81,304)*

sreynolds@daignaultiyer.com

ksprenger@daignaultiyer.com

djackson@daignaultiyer.com

smcbride@daignaultiyer.com

cbiyer@daignaultiyer.com

ehadi@daignaultiyer.com

lmeroueh@daignaultiyer.com

DAIGNAULT IYER LLP

8229 Boone Boulevard

Suite 450

Vienna, VA 22182

**Not admitted in Virginia*

Attorneys for SoundClear Technologies LLC

CERTIFICATE OF COMPLIANCE

Pursuant to 37 C.F.R. § 42.24(d), the undersigned certifies that this Patent Owner's Sur-reply complies with the type-volume limitations of 37 C.F.R. §§ 42.24(b) because it contains 4,665 words, as determined by the word-processing program used to prepare the brief, excluding the parts of the brief exempted by 37 C.F.R. § 42.24(a)(1).

Dated: April 21, 2026

Respectfully Submitted,

By: /Steven J. Reynolds/

Steven J. Reynolds
USPTO Reg. No. 61,445

CERTIFICATE OF SERVICE

The undersigned hereby certifies that the foregoing PATENT OWNER'S SUR-REPLY was served electronically via email on April 21, 2026, on the following counsel of record for Petitioner:

Colin B. Heideman
Jeremy A. Anapol
Christie R.W. Matthaei
Nathan D. Reeves
Logan P. Young
MORGAN, LEWIS & BOCKIUS LLP
colin.heideman@morganlewis.com
jeremy.anapol@morganlewis.com
christie.matthaei@morganlewis.com
nathan.reeves@morganlewis.com
logan.young@morganlewis.com
amazon-soundclear@morganlewis.com

Dated: April 21, 2026

Respectfully Submitted,

By: /Steven J. Reynolds/

Steven J. Reynolds
USPTO Reg. No. 61,445