

Filed on behalf of: Meta Platforms, Inc.

Filed: October 2, 2024

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.

Petitioner,

v.

JAWBONE INNOVATIONS, LLC,

Patent Owner.

Case No. IPR2023-01352

U.S. Patent No. 7,246,058

PETITIONER'S REPLY TO PATENT OWNER'S RESPONSE

TABLE OF CONTENTS

I.	INTRODUCTION.....	1
II.	CLAIM CONSTRUCTION	1
A.	“voicing sensor”.....	1
1.	The claims recite an acoustic throat microphone as the “voicing sensor.”	2
2.	The ’058 patent discloses acoustic-only embodiments.....	3
3.	PO’s Construction Lacks Merit.....	5
a.	PO’s construction is unsupported by the intrinsic record.....	5
b.	PO’s cited extrinsic evidence is irrelevant.....	7
c.	PO’s construction cannot be correct because it renders the claims invalid under §112.	8
B.	“physiological information”	9
C.	“cross-correlation data”	11
1.	PO’s narrow construction conflicts with the intrinsic evidence.	11
2.	Extrinsic evidence does not support PO’s construction.....	12
III.	GROUND 1: MEISEL IN VIEW OF KANEDA AND FREIDIN RENDER OBVIOUS CLAIMS 1-2 AND 4-5	14
A.	Meisel Discloses A “Voicing Sensor”	14
1.	Meisel discloses a voicing sensor under PO’s construction.	15
2.	PO’s arguments that Meisel does not disclose a “voicing sensor” under its construction are flawed.	16
B.	Meisel Discloses “Physiological Information” Even Under PO’s Incorrect Construction	18
C.	Meisel in View of Freidin Renders Obvious the Cross- Correlation Data Elements Recited in Claims 1 and 2	19

1.	PO applies the wrong legal standard for motivation to combine.	19
2.	Meisel combined with Freidin renders obvious the cross-correlation data elements even under PO's construction.	20
D.	Meisel in Further View of Kaneda Renders Obvious The Identifying Information of Acoustic Signals as Noise Limitation	24
IV.	GROUND 2-3: PO DOES NOT SEPARATELY ARGUE PATENTABILITY OF THESE CLAIMS.....	28
V.	CONCLUSION	28

TABLE OF AUTHORITIES

Page(s)

Cases

<i>AK Steel Corp. v. Sollac and Ugine</i> , 344 F.3d 1234 (Fed. Cir. 2003).....	8
<i>Baxalta Inc. v. Genentech, Inc.</i> , 972 F.3d 1341 (Fed. Cir. 2020).....	2
<i>Bayer Healthcare LLC v. Baxalta Inc.</i> , 989 F.3d 964 (Fed. Cir. 2021).....	6
<i>CCS Fitness, Inc. v. Brunswick Corp.</i> , 288 F.3d 1359 (Fed. Cir. 2002).....	7, 12, 13, 14
<i>Geneva Pharms., Inc. v. GlaxoSmithKline PLC</i> , 349 F.3d 1373 (Fed. Cir. 2003).....	9
<i>GPNE Corp. v. Apple Inc.</i> , 830 F.3d 1365 (Fed. Cir. 2016).....	1
<i>Intel Corp. v. PACT XPP Schweiz AG</i> , 61 F.4th 1373 (Fed. Cir. 2023).....	20
<i>Intel Corp. v. Qualcomm Inc.</i> , 21 F.4th 784 (Fed. Cir. 2021).....	21
<i>KSR Int'l Co. v. Teleflex Inc.</i> , 550 U.S. 398 (2007)	20
<i>Panduit Corp. v. Dennison Mfg. Co.</i> , 810 F.2d 1561 (Fed. Cir. 1987).....	2
<i>Ruckus Wireless, Inc. v. Innovative Wireless Sols., LLC</i> , 824 F.3d 999 (Fed. Cir. 2016).....	9
<i>Trustees of Columbia Univ. v. Symantec Corp.</i> , 811 F.3d 1359 (Fed. Cir. 2016).....	2
<i>Unwired Planet, LLC v. Google Inc.</i> , 841 F.3d 995 (Fed. Cir. 2016).....	20

<i>Vitronics Corp. v. Conceptronic, Inc.</i> , 90 F.3d 1576 (Fed. Cir. 1996).....	9
--	---

Other Authorities

37 C.F.R. §42.24.....	30
37 C.F.R. §42.24(c)	30
37 C.F.R. §42.24(c)(1)	30

EXHIBIT LIST

EX1001	U.S. Patent No. 7,246,058 (“the ’058 patent”)
EX1002	File History of U.S. Patent No. 7,246,058
EX1003	Declaration of Dr. David V. Anderson, Ph.D.
EX1004	CV of Dr. David V. Anderson, Ph.D.
EX1005	U.S. Patent No. 4,718,096 (“Meisel”)
EX1006	U.S. Patent No. 5,208,864 (“Kaneda”)
EX1007	U.S. Patent No. 6,148,313 (“Freidin”)
EX1008	U.S. Patent No. 4,239,936 (“Sakoe”)
EX1009	Signal Processing Toolbox for Use with MATLAB, User’s Guide, Version 4.2 (January 1999) (“MATLAB Manual”)
EX1010	<i>Reserved</i>
EX1011	Steven W. Smith, The Scientist and Engineer’s Guide to Digital Signal Processing (2d ed. 1999) (“Smith”)
EX1012	Richard G. Lyons, <i>Understanding Digital Signal Processing</i> (1996) (“Lyons”)
EX1013	U.S. Patent No. 5,724,485 (“Rainton”)
EX1014	U.S. Patent No. 5,774,847 (“Chu”)
EX1015	U.S. Patent No. 6,694,028 (“Matsuo”)
EX1016	Vinay K. Ingle and John G. Proakis, <i>Digital Signal Processing Using MATLAB</i> (2000) (“Ingle”)
EX1017	U.S. Patent No. 6,219,645 (“Byers”)
EX1018	U.S. Patent No. 6,246,760 (“Makino”)
EX1019	<i>Reserved</i>

EX1020	U.S. Patent No. 6,760,449 (“Matsuo 2”)
EX1021	U.S. Patent No. 6,766,019 (“Benesty”)
EX1022	<i>Reserved</i>
EX1023	Richard O. Duda and Peter E. Hart, <i>Pattern Classification and Scene Analysis</i> (1973) (“Duda”)
EX1024	David M. Green and John A. Swets, <i>Signal Detection Theory and Psychophysics</i> (1988) (“Green”)
EX1025	J. R. Deller et al., “ShortTerm Processing of Speech,” in <i>Discrete-Time Processing of Speech Signals</i> , IEEE (2000) (“Deller”)
EX1026	<i>Reserved</i>
EX1027	Declaration of Lauren Gluckman regarding Lyons
EX1028	Online Copyright Office Public Catalog Record of Lyons
EX1029	<i>Reserved</i>
EX1030	<i>Reserved</i>
EX1031	<i>Reserved</i>
EX1032	<i>Reserved</i>
EX1033	<i>Reserved</i>
EX1034	<i>Reserved</i>
EX1035	<i>Reserved</i>
EX1036	<i>Reserved</i>
EX1037	<i>Reserved</i>
EX1038	U.S. Provisional Application No. 60/368,209 (“the ’209 provisional”)

EX1039	Deposition Transcript of Akbar M. Sayeed, Ph.D.
EX1040	Gregory C. Burnett, <i>The Physiological Basis of Glottal Electromagnetic Micropower Sensors (GEMS) and Their Use in Defining an Excitation Functions for the Human Vocal Tract</i> , Thesis, University of California at Davis (1999)
EX1041	U.S. Patent No. 9,196,261 to Burnett et al.
EX1042	Reply Declaration of Dr. David V. Anderson, Ph.D.

I. INTRODUCTION

PO’s arguments are premised on unsupported, overly narrow claim constructions and misunderstandings of the prior art. But under any reasonable construction, as well as PO’s incorrect constructions, the prior art renders obvious the challenged claims.

II. CLAIM CONSTRUCTION

A. “voicing sensor”

Petitioner’s Proposed Construction	PO’s Proposed Construction
plain and ordinary meaning	“a sensor that is able to detect human tissue motion associated with speech and is unaffected by environmental noise”

The Petition established that Meisel teaches the same “voicing sensor” embodiment recited in a dependent claim of the ’058 patent—an acoustic throat microphone. Pet. §V.A.4.c. Accordingly, the Board need not provide an express construction. Rather, the Board need only resolve the dispute on this term—whether a “voicing sensor” includes an acoustic throat microphone. *GPNE Corp. v. Apple Inc.*, 830 F.3d 1365, 1373-74 (Fed. Cir. 2016). The claims and specification confirm it does.

In a transparent attempt to evade the prior art, PO proposes a narrow construction that it further construes to limit “voicing sensor” to non-acoustic sensor. Not only is this contrary to the explicitly claimed “*acoustic* throat microphone,” but

it is also inconsistent with the specification. PO's construction must be rejected. *See Baxalta Inc. v. Genentech, Inc.*, 972 F.3d 1341, 1346 (Fed. Cir. 2020) (rejecting construction that excludes "explicitly claimed embodiments" recited in dependent claims).

1. The claims recite an acoustic throat microphone as the "voicing sensor."

Claim 1 of the '058 patent recites a "voicing sensor that receives physiological information associated with human voicing activity," identifying the voicing sensor only by its function. EX1001, claim 1. Claim 5 expressly recites that "the voicing sensor includes at least one detector selected from a group including...acoustic throat microphones," without the further qualifications that PO seeks to add. *Trustees of Columbia Univ. v. Symantec Corp.*, 811 F.3d 1359, 1370 (Fed. Cir. 2016).

PO argues lexicography, stating the "'058 patent *explicitly defines* the voicing sensor as a *non-acoustic* sensor." POR 8 (emphasis added); *Panduit Corp. v. Dennison Mfg. Co.*, 810 F.2d 1561, 1576 (Fed. Cir. 1987). But PO cannot reconcile its construction with the "*acoustic* throat microphone[]" recited in claim 5. EX1042, ¶¶20-21. PO asserts that a non-acoustic sensor is required because "[a] POSITA would understand that a throat microphone is (and must be) unaffected by environmental noise, by touching the skin of a user to sense

vibration.” POR 11. But as PO’s own expert admits, a non-acoustic sensor can be affected by non-acoustic noise. EX1039, 30:7-15, 32:16-34:11. Thus, PO’s construction of its own proposed construction that “unaffected by environmental noise” necessarily requires a non-acoustic sensor (POR 10) is faulty.

PO also argues that the claim language supports its assertion that the “voicing sensor” must be non-acoustic because it draws a distinction between “acoustic signals” and “physiological information.” POR 13. Incorrect. The claim language requires “the physiological information [received by the voicing sensor] and an acoustic signal *received at one of the two microphones.*” A POSITA would have understood that this language does not preclude the physiological information from being an acoustic signal. EX1042, ¶¶28-29. To the contrary, read in light of the specification and other claims, a POSITA would have understood that the voicing sensor could receive acoustic signals, non-acoustic signals, or both, including from an acoustic throat microphone. *Id.*, ¶29.

2. The '058 patent discloses acoustic-only embodiments.

The specification discloses two types of speech recognition systems for detecting voiced speech: NAVSAD (Non-Acoustic Sensor Voiced Speech Activity Detection) and PSAD (Pathfinder Speech Activity Detection). EX1001, 2:20-24; *see also* EX1042 ¶15. The PSAD “supports an acoustic approach that functions in most low- to medium-noise environments,” while the NAVSAD system “adds a non-

acoustic sensor to enable detection of voiced speech in any environment.” *Id.*, 3:10-15.

The “PSAD system couples [acoustic] microphones 10” (*id.*, 2:41-43) to a processor and can be used “to detect both voiced and unvoiced speech” (*id.*, 9:44-46). PSAD may “operate independently” of the NAVSAD, or as the sole voiced speech detection. *Id.*, 2:48-49, 3:12-13. Figure 2 shows the acoustic PSAD system being used to determine voicing activity, particularly “in low acoustic noise environments” where it is “highly sensitive.” *Id.*, 2:45-49.

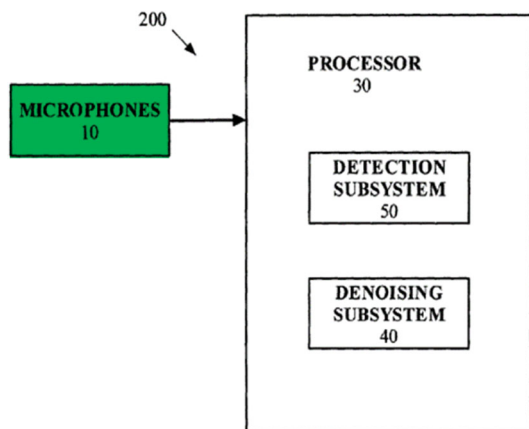


Figure 2

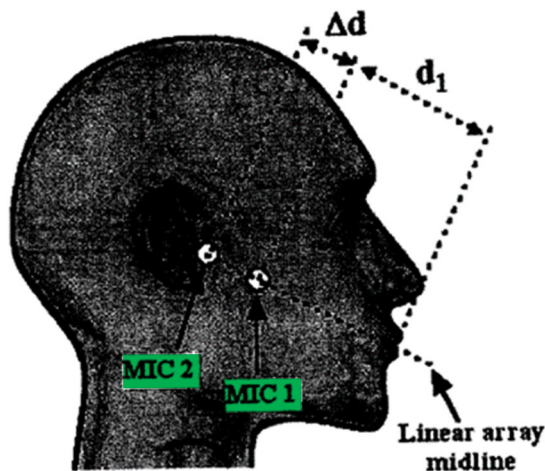


Figure 7

EX1001, FIGS. 2, 7 (annotated).

The '209 provisional, incorporated by reference in the '058 patent, confirms that “voicing sensor” includes acoustic throat microphones. EX1042 ¶16. The '209 provisional explains that “[t]he external VAD [(voicing activity detectors)] device

may include...acoustic microphones and associated software,” that “[t]wo-microphone acoustic VADs” can “operate without any external hardware, just using two microphones,” and that the “Pathfinder VAD” can “determine when voicing is occurring” “with any of the microphone configurations above with no modifications.” EX1038, 8, 15, 16. Thus, a VAD including an acoustic microphone is “a voicing sensor that receives physiological information associated with human voicing activity” as expressly recited in the claim.

3. PO's Construction Lacks Merit

a. PO's construction is unsupported by the intrinsic record.

PO cherry picks limitations untied to a “voicing sensor” to improperly narrow the term. PO argues its construction “comes directly from the specification,” citing to the '058 patent, 3:61-63. But this sentence is no definition. *See* EX1042, ¶19. First, “voicing sensor” is never referenced in the same paragraph, or even adjacent paragraphs. Second, the sentence refers to “alternative embodiments,” which are among many embodiments (including acoustic embodiments) discussed in the specification. Third, that alternative embodiments “may use” such a sensor has nothing to do with whether the claimed “voicing sensor” should be so limited.

Indeed, the intrinsic record contradicts PO's strained narrowing, broadly stating “[a]ny conventional acoustic method may also be used with either or both

microphones...to construct the VAD signal.” EX1038, 11. There is no explicit definition nor any indication that the patentee clearly and unmistakably disclaimed the scope of “voicing sensor” to non-acoustic sensors, as PO contends. *Bayer Healthcare LLC v. Baxalta Inc.*, 989 F.3d 964, 975 (Fed. Cir. 2021).

PO also argues “every embodiment in the patent is unaffected by environmental noise” because the “preferred embodiment...uses the General Electromagnetic Movement Sensor (GEMS),” and thus the claimed “voicing sensor must be unaffected by environmental noise.” POR 8. But GEMS is affected by environmental noise. PO’s expert admitted that noise can be non-acoustic, EX1039, 30:7-15, 32:16-34:11, and GEMS is affected by non-acoustic noise caused by movement near the vocal cord during speech. EX1042, ¶18. The ’058 patent acknowledges this, explaining that “[a]nything that moves near the [GEMS] sensor will cause a change in phase of the reflected wave that will be amplified and displayed as a change in voltage output from the sensor” and appear as noise signals. EX1001, 4:6-9; EX1042, ¶18. The inventor’s own thesis further confirms that GEMS is affected by substantial environmental noise, including “the jaw and tongue moving, the blood coursing through arteries and veins, the vertical motion of glottis as speech begins and ends, and any motion of the GEMS itself with respect to the neck and trachea.” EX1040, 49; EX1042, ¶18. Regardless, even if GEMS were

entirely unaffected by environmental noise, there is no disclaimer that justifies limiting the voicing sensor to this single preferred embodiment. *Supra* §II.A.2.

PO relies on the patent to argue that a “non-acoustic sensor” is “required to ensure good performance.” POR 10. But when read in its entirety, this disclosure states that a non-acoustic sensor *may* be required in “occasional[]” situations where “noise levels [are] high enough that the speech will be nearly undetectable.” EX1001, 4:53-57. This merely confirms that *one* type of VAD may be non-acoustic and does not exclude the acoustic VADs. *See CCS Fitness, Inc. v. Brunswick Corp.*, 288 F.3d 1359, 1366 (Fed. Cir. 2002).

PO further argues that an acoustic throat microphone is the claimed voicing sensor only if it “touch[es] the skin of a user.” POR 11. But this argument is unsupported. Indeed, claim 5 simply recites that the voicing sensor includes an “acoustic throat microphone[s]” without any qualification, and the specification never describes the microphone being in contact with the speaker's skin. *Supra* §§II.A.1-II.A.2. And PO's own expert confirmed that a sensor need not touch human tissue to detect its motion. EX1039, 40:1-12; *see also* 38:8-25; EX1042 ¶17.

b. PO's cited extrinsic evidence is irrelevant.

PO argues that throat microphones must be skin surface microphones by pointing to definitions from Merriam Webster and a 2004 article comparing “throat microphones” to “conventional acoustic microphones.” POR 11-12. But these

references are irrelevant because none address the actual embodiment—*acoustic* throat microphones. EX1042 ¶21. By its very name, an *acoustic* throat microphone detects acoustic signals. *Id.* PO's attempt to read out "acoustic" from "acoustic throat microphones" must fail.

c. PO's construction cannot be correct because it renders the claims invalid under §112.

PO's proposed construction should also be rejected because it renders the claims invalid. *AK Steel Corp. v. Sollac and Ugine*, 344 F.3d 1234, 1244 (Fed. Cir. 2003). The specification does not explain the meaning of "environmental noise" nor how to determine when a voicing sensor is deemed "unaffected" by it. Injecting these requirements into this claim term would render the claims non-enabled because no sensor would meet that construction.

PO argues that the "acoustic throat microphone" in claim 5 "does not remove the requirement to be unaffected by environmental noise." POR 11. But PO ignores that the specification discloses acoustic non-contact microphones as voicing sensors. *Supra* §II.A.2. PO's alleged "requirement" is merely one NAVSAD "embodiment" and is not coextensive with the claim scope. PO's attempt to narrowly construe "acoustic throat microphone" in claim 5 to reconcile its "voicing sensor" construction cannot be correct because the narrowing lacks written description.

Ruckus Wireless, Inc. v. Innovative Wireless Sols., LLC, 824 F.3d 999, 1004 (Fed. Cir. 2016).

Adopting PO’s proposed construction would also exclude the preferred non-acoustic GEMS embodiment that PO relies on, as explained in §II.A.3.a. EX1001, 3:57-60. *See Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1583-84 (Fed. Cir. 1996).

PO’s construction would further render the claims indefinite because the specification fails to provide any guidance to determine the outer boundaries of “unaffected,” such that a POSITA could not determine when a sensor would be deemed “unaffected by environmental noise.” *See Geneva Pharms., Inc. v. GlaxoSmithKline PLC*, 349 F.3d 1373, 1384 (Fed. Cir. 2003).

B. “physiological information”

Petitioner’s Proposed Construction	PO’s Proposed Construction
plain and ordinary meaning	“information on human tissue motion”

There is nothing in the intrinsic record to narrow the scope of “physiological information” from its plain and ordinary meaning. EX1042 ¶¶22-30. Thus, a POSITA would understand “physiological” to mean more than “human tissue motion.”

As PO's own expert admitted: "physiology is much bigger than human tissue motion." EX1039, 48:10-18; *see also* EX1042 ¶23. Indeed, another patent by the same inventor confirms a VAD that "uses physiological information to determine when a speaker is speaking" can use "an acoustic microphone that is being used to detect acoustic frequency signals that correspond to the user's speech directly from the skin of the user." EX1040, 3:41-50. PO's expert explained "[s]ounds are changes in air pressure or air molecules; vibrating, if you will...and typical microphone, it's diaphragm, senses those changes in pressure." EX1039, 38:11-14. Notably, Dr. Sayeed admitted that "at a very basic level, you can argue that, yes, it does" when asked whether "an acoustic signal produced by a person speaking contain[s] physiological information." EX1039, 50:24-51:16.

Like "voicing sensor," PO argues that its construction "comes directly from the specification." POR 14. But like "voicing sensor," the cited passages are untied to the term "physiological information." In fact, the term only appears in the Abstract as part of a discussion of certain embodiments (among many embodiments) of the system: "Further, embodiments of the system include non-acoustic sensors that receive physiological information to aid in identifying voiced speech." EX1001, Abstract. There is no mention of human tissue motion in the Abstract, and a POSITA reading this passage would not understand that "physiological information" is so limited. EX1042, ¶26. Neither the Abstract nor

rest of the specification limit “physiological information,” let alone to human tissue motion.

C. “cross-correlation data”

Petitioner's Proposed Construction	PO's Proposed Construction
plain and ordinary meaning	“data generated by multiplying and adding the result of a windowed set of samples at one time from one source by a windowed set of samples from another source where the two sets of samples may be shifted in time relative to each other”

PO argues that the '058 patent uses “cross-correlation” “in its typical meaning as being generated by multiplying two samples together, with those samples being at the same or a different time and adding up this series of multiplications over a finite time window.” POR 17. But PO's construction is not the “typical meaning,” and does not identify anything in the intrinsic record that clearly and unmistakably shows that the patentee acted as a lexicographer to limit the scope of the term to this subset of cross-correlation data. *CCS Fitness*, 288 F.3d at 1366.

Ultimately, PO argues that because *some* extrinsic references disclose a specific cross-correlation function, then *all* cross-correlation data must be generated according to that function. POR 16. That is a logical fallacy. No such definition exists in the intrinsic record, and PO cites to none.

1. PO's narrow construction conflicts with the intrinsic

evidence.

Claim 1 recites generating “cross-correlation data” “between physiological information and an acoustic signal received at one of the two microphones.” EX1001, claim 1. There is no recitation of multiplication, adding, windowing, sampling, or time-shifting to limit the term's meaning. EX1001, claim 1.

The '058 patent does not define “cross-correlation.” Instead, it states only that “cross-correlation” is calculated “between the acoustic signal from microphone 1 and the sensor data” “[t]o further differentiate sensor noise from tissue motion.” EX1001, 5:2-3, 5:11-12. PO's narrow construction is thus unsupported by the intrinsic record and is also inconsistent with how a POSITA would have understood “cross-correlation.” EX1003, ¶¶53-75; EX1042, ¶¶31-37; Pet. §§V.A.4.e-f.

2. Extrinsic evidence does not support PO's construction.

PO argues its construction is correct because similar examples are found in “multiple exhibits” of Petitioner. POR 16. PO conflates the issue. It is immaterial that there are multiple extrinsic references describing one example of cross-correlation because the intrinsic record does not limit cross-correlation data to that single example. *CCS Fitness*, 288 F.3d at 1366. PO's construction is inconsistent with the intrinsic record and a POSITA's understanding of the term. Pet. §§V.A.4.e-f; EX1003, ¶¶53-75; EX1042, ¶¶35-36.

PO argues that cross-correlation “requires the ability to shift the samples in time.” POR 17. But the specification’s description that there can be “delays” in the cross-correlation function does not mean that the patentee clearly and unmistakably narrowed cross-correlation data to PO’s specific mathematical formula. PO’s construction, for example, requires multiplying the signals, but the ’058 patent does not describe multiplying the samples, let alone clearly and unambiguously identify such multiplication as required.¹ See *CCS Fitness*, 288 F.3d at 1366.

Furthermore, even if PO’s “shift[ing] the samples in time” is required for cross-correlation, Meisel discloses it. Meisel adjusts its input speech signals “to follow the envelope of the speech signal” and the “output of each of the amplitude detectors 501 and 502 thus represents the amplitude of the input signal on an instant-by-instant basis,” which a POSITA would have understood refers to shifting signals to evaluate the same speech segment (window). EX1005, 9:6-11; EX1042, ¶37. A POSITA would have recognized the need to adjust Meisel’s signals to account for delays in its multi-microphone speech recognition system because different

¹ PO’s construction also requires the data is sampled, even though its expert admitted that cross-correlation can be in the analog domain (EX1039, 63:8-11), and windowed, even though PO’s expert admitted the equations in his declaration are not windowed and contradict PO’s construction (*id.*, 65:16-23).

microphones positioned at different locations would receive speech signals at different times. *Id.* Thus, like the '058 patent, Meisel extracts a windowed signal (“envelope”) over a period of time to correlate different signals that may be misaligned in time by several milliseconds. *Id.*

PO misrepresents Dr. Anderson's testimony, arguing that he “further confirmed that a POSITA would understand that ‘cross-correlation’ requires multiplication of samples.” POR 18. Dr. Anderson testified that “there are multiple functions called ‘cross-correlation,’” but not all of them involve multiplying signals. EX2001, 28:18-30:6. Thus, cross-correlation data includes functions that do not involve multiplication. *Id.*

III. GROUND 1: MEISEL IN VIEW OF KANEDA AND FREIDIN RENDER OBVIOUS CLAIMS 1-2 AND 4-5

A. Meisel Discloses A “Voicing Sensor”

Meisel discloses the “voicing sensor” elements because its throat microphone detects physiological information associated with human voicing activity. *See* Pet. §§V.A.4.c, V.A.5.b.

PO's assertion that Meisel does not disclose a “voicing sensor” relies solely on its improper claim construction. POR 20. When the claims are properly construed to include acoustic microphones, *supra* §II.A, PO does not dispute that Meisel discloses this claim element.

1. Meisel discloses a voicing sensor under PO's construction.

Even if PO's construction is adopted, Meisel's throat microphone is a voicing sensor. As an initial matter, the POR observes that Meisel's throat microphone is unaffected by environmental noise because "the placement of the microphones will prevent noise." POR 34.

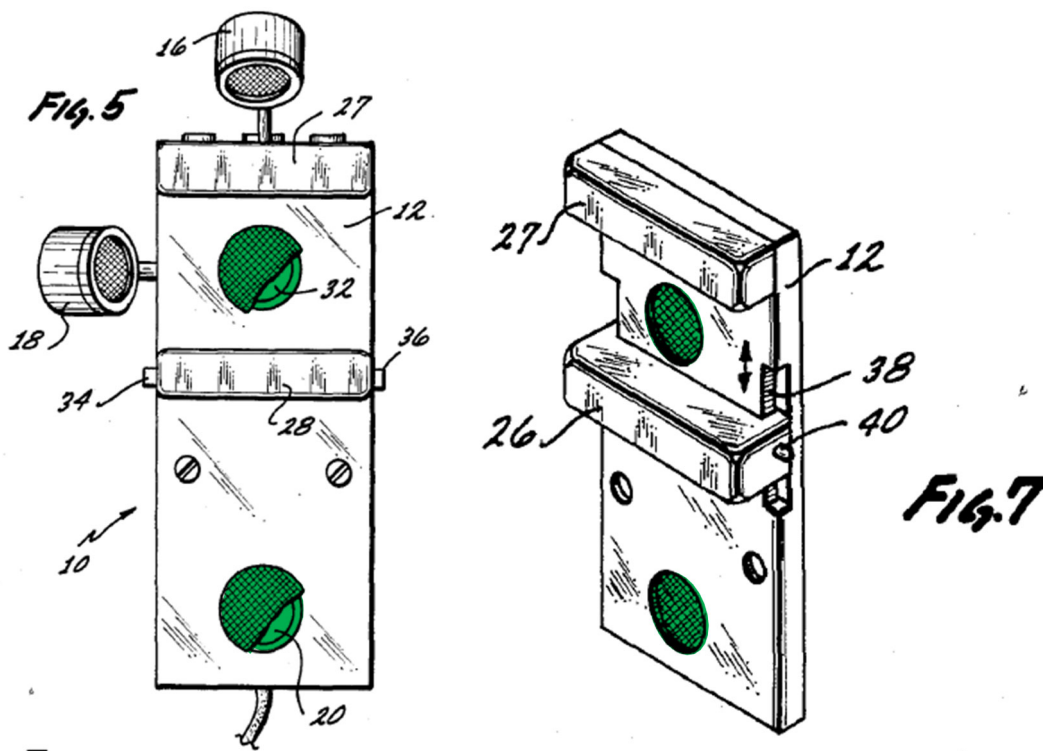
Meisel's acoustic throat microphone is a sensor that receives physiological information in the form of human tissue motion associated with speech. Meisel's "throat microphone...*detects the vibration of the vocal cords*, or 'voicing,'" and PO and its expert do not dispute that detecting the vibration of the vocal cords is receiving human tissue motion associated with speech. EX1005, 2:33-35 (emphasis added); EX2001, 49:4-23; EX1039, 38:11-17, 41:18-20; *accord* EX1001, 3:55-57.

Meisel's throat microphone is also protected from environmental noise and is a "close-talking microphone[]" that "*respond[s] very little to background noise.*" EX1005, 2:58-60 (emphasis added). Meisel's speech recognition system can be used "under conditions where there is background noise and where the speaker can continue to perform other duties." *Id.*, 1:10-15. Further, Meisel teaches that its "microphones are supported in a solid sound absorbing substance" that prevents sound "from entering the microphone except through desired external ports." *Id.*, 8:9-21. Thus, Meisel discloses the claimed voicing sensor even under PO's proposed construction.

2. PO's arguments that Meisel does not disclose a "voicing sensor" under its construction are flawed.

PO argues that Meisel's throat microphone is not a "voicing sensor" because it is affected by environmental noise. POR 25. Specifically, PO argues that Meisel's housing "would allow for substantial environmental noise to reach the throat microphone" because of "spacing between the throat microphone and the user's skin" and "the openings to both sides and below the microphone." *Id.* But PO's argument is contradicted by the '209 provisional, which states that "close-talk microphones," like Meisel's (EX1005, 2:54-60), achieve their "best performance" when there is a gap between the speaker and the microphone. EX1038, 10-11 (when "'close-talk' microphones (which de-emphasizes acoustic sources more than a few centimeters away from the microphone)" are used, "[t]he best performance is seen when the two microphones are quite close (less than approximately 5 cm), but when the [unidirectional microphone] is far enough away from the user's mouth (approximately 10-15 centimeters depending on the microphone) so that the UNI directionality functions efficiently"). Further, the '209 provisional does not state that the close-talk microphone must be closed on the sides and below. *See id.* Thus, PO's own disclosure confirms that Meisel's unidirectional throat microphone is at its optimum when placed with a gap, as Meisel confirms that the microphone "respond[s] very little to background noise." EX1005, 2:58-60.

Further, Meisel's microphones are embedded in housing, having a ring around where the throat microphone that is "a solid sound absorbing substance" that prevents sound "from entering the microphones" and "attenuates sound vibrations in the housing." EX1005, 8:9-21. Figure 7 shows where the acoustic throat microphone is recessed to help protect the microphone from sounds emanating from different directions. EX1042, ¶39. Thus, Meisel's microphones are insulated from, and therefore unaffected by, environmental noise. *Id.*



EX1005, FIGS. 7 (annotated).

PO argues that an acoustic throat microphone satisfies the voicing sensor limitation only when it touches the user's skin. POR 12, 22. But this relies on PO's

improper construction, which is contradicted by the intrinsic record. *Supra* §§II.A.2-II.A.3. None of the claims contain any limitation on the construction or placement of the acoustic throat microphone, and the specification never describes, let alone requires, skin contact. *Supra* §§II.A.2-II.A.3. Indeed, the '209 provisional explains that acoustic microphones perform best “when the [unidirectional microphone] is far enough away from the user’s mouth (approximately 10-15 centimeters depending on the microphone)” (EX1038, 10-11), indicating that an acoustic throat microphone should not contact the speaker’s skin. EX1042, ¶¶40-41.

B. Meisel Discloses “Physiological Information” Even Under PO’s Incorrect Construction

PO argues that Meisel’s throat microphone does not receive physiological information in the form of “human tissue motion.” POR 23. Incorrect. *Supra* §III.A.1. As PO itself admits, “Meisel’s throat microphone is described as a conventional microphone pointed at a speaker’s throat as depicted above, and is described as receiving *sounds* (*i.e.*, acoustic signals).” *Id.* A POSITA would understand that those sounds correspond to human tissue motion, particularly since this describes the function of the “acoustic throat microphones” recited in claim 5. EX1042, ¶¶42-45.

A POSITA would have understood that a sensor need not touch human tissue to detect its movement. *Id.*, ¶44. Instead, as PO’s expert admitted, the sensor, such

as a microphone, can detect vibrations traveling through air or solids, caused by human tissue motion. EX1039, 38:8-25.

Thus, even when the term is incorrectly narrowed to human tissue motion, Meisel discloses “physiological information.”

C. Meisel in View of Freidin Renders Obvious the Cross-Correlation Data Elements Recited in Claims 1 and 2

Even if the Board determines that Meisel does not disclose the claimed “cross-correlation data,” it is undisputed that Freidin discloses it. EX1042., ¶33. PO attempts to argue that there is no motivation to combine, but PO's arguments lack merit.

1. PO applies the wrong legal standard for motivation to combine.

PO argues that “Meisel already includes all the necessary functionality to determine the classification of a sound” and that there is no indication that “a POSITA would perceive Meisel's speech detection as lacking accuracy or even being susceptible to an improvement in accuracy.” POR 29. PO is wrong. There is no requirement to show a “deficiency” in the prior art reference to establish obviousness of a combination. To the contrary, a patent challenger “*does not need to show* that there was a known problem with the prior art system in order to articulate the required rational underpinning for the proposed combination.” *Unwired Planet, LLC v. Google Inc.*, 841 F.3d 995, 1002-03 (Fed. Cir. 2016)

(emphasis added); *see also Intel Corp. v. PACT XPP Schweiz AG*, 61 F.4th 1373, 1381 (Fed. Cir. 2023). PO's argument thus turns on the wrong legal framework.

PO's next assertion that obviousness cannot be shown because "*Meisel* does not state or identify a need" to modify itself, POR 29, was rejected 17 years ago by the Supreme Court. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 419 (2007) (rejecting obviousness test that would require a teaching, suggestion, or motivation in the prior art reference itself); EX1042, 17:18-18:3. It does not matter how well *Meisel* performed its disclosed functions, a POSITA would have been motivated to look for ways to improve the speech recognition system for more accurate speech recognition, including using cross-correlation data. Pet. §V.A.4.e; *see Intel Corp. v. Qualcomm Inc.*, 21 F.4th 784, 800 (Fed. Cir. 2021).

2. Meisel combined with Freidin renders obvious the cross-correlation data elements even under PO's construction.

Even under PO's flawed construction, *Meisel* in view of *Freidin* would have rendered obvious the cross-correlation data elements of claims 1 and 2. *See* Pet. §§V.A.4.e, V.A.5.c. PO does not dispute that *Freidin* discloses "cross-correlation

data” under its construction. *See generally* POR.² PO's motivation to combine argument (POR 32) lacks merit.

PO acknowledges that Meisel's speech recognition system uses “a simple compare operation between two inputs” to determine similarities and differences between signals. POR 18. It would have been obvious to a POSITA to implement cross-correlation (as construed by PO)—another well-known signal comparison technique—to improve Meisel's speech recognition system to compare signals to distinguish voiced from unvoiced speech. *See* Pet. §§V.A.4.e.i-ii.

A POSITA would have recognized that Freidin's cross-correlation is compatible with Meisel's system because it was well known that “XCORR is more susceptible to errors from acoustic noise, and in high (<0 dB SNR) environments is almost useless,” and Meisel also functions in environments with little “background noise, such as may occur in an office environment,” where functions like XCORR perform best. EX1001, 5:21-25; EX1005, 2:58-60; Pet. §V.A.4.e. Applying cross-correlation data, as construed by PO and disclosed in Freidin, to Meisel's system

² PO's expert admits a POSITA would have known how to obtain cross-correlation data. EX1039, 67:7-68:14. To the extent that PO argues Meisel's analog system is incompatible with Freidin's digital system, PO's expert admits analog-to-digital converters were known before 2001. *Id.*, 73:9-12.

would have involved nothing more than using a known technique (cross-correlation) to improve similar devices (Meisel's system applying a simple comparison of two signals) in the same way (signal processing for speech detection/recognition). EX1042, ¶28; Pet. §§V.A.4.e.i-ii. It would have involved only combining prior art elements (cross-correlation and speech detection/recognition) according to known methods (signal processing for speech detection/recognition) to yield predictable results (more accurate speech detection/recognition). *Id.*

PO's contrary arguments lack merit. PO argues that "[d]etecting similarities goes against the actual inventive concept in Meisel, which is to classify speech by looking at imbalances between microphone signals." POR 30. PO assumes that cross-correlation only detects similarities and Meisel only detects differences between two signals. PO is wrong. Meisel compares two signals, which necessarily includes determining similarities and differences between the signals. *See* Pet. §V.A.4.f. This is also true for cross-correlation. If the cross-correlation between the signals is greater than a predetermined threshold, the signals are more similar than different, and vice versa. *Id.*; EX1042, ¶53.

PO argues that "Meisel makes a comparison between the throat microphone and the mouth microphone without checking to see if anything exceeds a threshold." POR 32. But PO fails to address Petitioner's argument that a POSITA would have

found it obvious to implement Freidin's use of a predetermined cross-correlation threshold for a more accurate determination of voiced speech. Pet. §V.A.4.f.

Instead, PO merely argues that a POSITA would not have been motivated to combine Meisel's speech recognition system with Freidin's cross-correlation because "Meisel outputs a 1 from ADCI 58 when a sound is voiced, in which case the throat microphone amplitude is greater than the side mouth microphone amplitude." POR 28. Incorrect. Meisel's system recognizes speech as a voiced sound where the throat microphone and the side mouth microphone both have high signal amplitudes, so long as the throat microphone signal is sufficiently higher than the side mouth microphone signal. Pet. 29; EX1003, ¶¶58-59; EX1042, ¶49.

PO also argues that "no portion of Meisel uses, or could use, data suggesting that the throat microphone and mouth microphone both have high amplitudes" because Meisel would have a low cross-correlation when it recognizes a "voiced consonant" using ADCI 58. POR 31, 33. PO's analysis is incomplete and incorrect because it looks only at one edge case for detecting only *one* type of voiced speech—voiced stop consonants. EX1042, ¶50. A POSITA would have understood that Meisel's detection of vowels, which are also voiced speech, produces high cross-correlation values. Pet. §V.A.4.f; EX1005, 2:35-36, 12:25-33. Meisel discloses that vowels are "relatively easy to distinguish" due to the large signal amplitudes at both throat and side mouth microphones. EX1005, 15:49-51; Pet.

§V.A.4.f; EX1039, 12:21-25. A POSITA would have found it obvious to implement Freidin's cross-correlation data in Meisel's system to detect vowels, where a high cross-correlation value indicates that a vowel is being spoken. EX1042, ¶50; Pet. §V.A.4.f.

PO's argument also rings hollow because the '058 patent itself does not produce a high cross-correlation value between a mouth microphone signal and a GEMS signal for voiced stop consonants. EX1042, ¶51. The mouth microphone of the '058 patent would have a low signal amplitude and GEMS will have a high signal amplitude, leading to a low cross-correlation value. *Id.* PO's focus on voiced stop consonants leads to an absurd result where the disclosed speech recognition system itself would not practice the claim.

D. Meisel in Further View of Kaneda Renders Obvious The Identifying Information of Acoustic Signals as Noise Limitation

Meisel in further view of Kaneda discloses identifying "information of the acoustic signals as noise when the difference parameters are less than the gain threshold" recited in claims 1 and 2. Pet. §§V.A.4.i, V.A.5.g.

PO does not dispute that Kaneda discloses this limitation. Pet. §V.A.4.i; *see* POR 34-37. Instead, PO argues that a POSITA would not have been motivated to combine Meisel with Kaneda because Meisel does not disclose a noise detection process. POR 37. PO is wrong.

Meisel discloses noise detection because its system is designed for use “in the normal work environment of the speaker, particularly under conditions where there is background noise.” EX1042, ¶53. Meisel’s system generates a difference between two microphone signals (A-B) and compares the difference to a threshold (T1); and if the difference is greater than T1, recognizes the signals as speech. Pet. §V.A.4.h. If the difference is less than T1, Meisel’s system recognizes the signals as noise. *Id.*; EX1042, ¶54.

PO argues that Meisel does not identify noise because “the placement of the microphones will prevent noise,” eliminating the need for noise detection. POR 34. But a POSITA would have been motivated to add noise detection to Meisel to further refine its speech signals and improve its accuracy. Pet. §V.A.4.i. PO also argues that, if Meisel discloses “noise cancellation,” “it further supports that Meisel’s throat microphone is not a voicing sensor as claimed.” POR 34. First, PO’s claim construction of “voicing sensor” is wrong. *Supra* §II.A. Second, even if PO’s construction is adopted, a POSITA would not have understood it to require the voicing sensor to be completely isolated from noise for the reasons explained in Section II.A. Instead, a POSITA would have understood that some level of noise is permitted during speech recognition. EX1042, ¶¶55-56.

PO also argues that Meisel does not detect noise because it uses a “Boolean variable and difference to determine if a plosive or fricative is being pronounced.”

POR 36. PO misunderstands Meisel. Meisel's system uses two gain thresholds, T1 and T2. Pet. §V.A.4.h. The T1 threshold is set to filter signals associated with background noise before proceeding to the next step. *Id.* It would have been well within a POSITA's skill to set T1 appropriately to filter low level background noise while leaving speech signals intact. *Id.*; EX1003, ¶¶83-88 EX1042, ¶55. For example, each ADCI is designed to pick up signals with different amplitudes, and thus, a POSITA would have set different T1 for each ADCI to accurately distinguish noise and speech. *Id.* The speech signals then proceed to the next comparison against T2, which is designed to detect different types of speech, for example, voiced or unvoiced speech. Pet. §V.A.4.h. Thus, each ADCI unit detects noise at the T1 step before it outputs a Boolean variable. *Id.*

PO argues that Meisel "faults Sakoe for using one of its two microphones for non-speech purposes." POR 34. PO again misses the mark. Sakoe's purported shortcoming is precisely why a POSITA would have been motivated to add another speech microphone to achieve more accurate speech recognition. Pet. §V.A.4.i; EX1042, ¶56. Meisel explains that, in Sakoe, "a single microphone yields no direct clue as to the physiological source (e.g., mouth, nose, or throat) of the sound, even though this information is valuable in recognition." EX1005, 1:58-61. This would have motivated a POSITA to combine Meisel and Kaneda to solve this known problem and improve the accuracy of speech recognition. Pet. §V.A.4.i.i-ii.

PO argues that, “even if the Kaneda process were added in parallel to the Meisel system, it would result in signals that Meisel considers speech to be classified as noise.” POR 36. Not true. Kaneda's noise detection system would be added to the T1 step before Meisel's system outputs a Boolean variable. Pet. §V.A.4.i. Thus, noise signals that fall below T1 would not proceed to the next step of identifying the part of speech, and thus, an ADCI would not output any Boolean variable for those signals. *Id.*

PO argues that “under Petitioner's proposed system, if there are similar amplitude signals from two microphones, then that portion of the entire output would be classified as noise.” POR 37. False. Meisel would classify signals as noise only if two signals have sufficiently similar amplitudes (i.e., $A-B$ is 0 or less than T1), or if signal B is greater than signal A. Pet. §V.A.4.i. A POSITA would have understood that, during speech, microphones at different locations, for example, front mouth microphone (signal A) and side mouth (signal B) microphone, would have different speech signal amplitudes because they are positioned differently with respect to the sound source. Pet. §V.A.4.i; EX1042, ¶57. Furthermore, during speech, the side mouth microphone would have lower signal amplitudes because it is at an angle with respect to the speaker's mouth, unlike the front mouth microphone. *Id.* If these criteria are not met, Meisel properly determines the signals as noise. *See* Pet. §V.A.4.i.

**IV. GROUND 2-3: PO DOES NOT SEPARATELY ARGUE
PATENTABILITY OF THESE CLAIMS**

PO does not provide any opposition to Grounds 2 and 3 other than to defer to its opposition for Ground 1. POR 38. Thus, claims 3 and 4 are obvious as established in the Petition and above. Pet. §§V.B-V.C.

V. CONCLUSION

The Board should find all challenged claims unpatentable.

Dated: October 2, 2024

By: / Lisa K. Nguyen /
Lisa K. Nguyen (Reg. No. 58,018)
PAUL HASTINGS LLP
1117 S. California Avenue
Palo Alto, CA 94304
Telephone: 650-320-1800
lisanguyen@paulhastings.com

*Counsel for Petitioner
Meta Platforms, Inc.*

CERTIFICATE OF COMPLIANCE WITH 37 C.F.R. §42.24

I hereby certify that this Reply complies with the word count limitation of 37 C.F.R. §42.24(c)(1) because the Reply contains a total of 5,590 words, which is the sum calculated by Microsoft Word's word-count feature. This total excludes the cover page, signature block, and the parts of the Reply exempted by 37 C.F.R. §42.24(c).

Dated: October 2, 2024

By: / Lisa K. Nguyen /
Lisa K. Nguyen (Reg. No. 58,018)
PAUL HASTINGS LLP
1117 S. California Avenue
Palo Alto, CA 94304
Telephone: 650-320-1800
lisanguyen@paulhastings.com

*Counsel for Petitioner
Meta Platforms, Inc.*

CERTIFICATE OF SERVICE

The undersigned certifies that a true and correct copy of **PETITIONER'S REPLY TO PATENT OWNER'S RESPONSE** was served via electronic email on Patent Owner's counsel of record listed below, pursuant to its Mandatory Notices, on October 2, 2024:

Alfred R. Fabricant
Peter Lambrianakos
Vincent J. Rubino, III
Jacob Ostling
Richard Cowell
Enrique W. Iturralde
FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Telephone: (212) 257-5797
Email: ffabricant@fabricantllp.com
Email: plambrianakos@fabricantllp.com
Email: vrubino@fabricantllp.com
Email: jostling@fabricantllp.com
Email: rcowell@fabricantllp.com
Email: eiturralde@fabricantllp.com

By: / Lisa K. Nguyen /
Lisa K. Nguyen (Reg. No. 58,018)
PAUL HASTINGS LLP
1117 S. California Avenue
Palo Alto, CA 94304
Telephone: 650-320-1800
lisanguyen@paulhastings.com

*Counsel for Petitioner
Meta Platforms, Inc.*