

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SONOS, INC.,
Petitioner,

v.

GOOGLE LLC,
Patent Owner.

IPR2023-00119
Patent 10,593,330 B2

Before KEVIN F. TURNER, BARRY L. GROSSMAN, and
TERRENCE W. McMILLIN, *Administrative Patent Judges*.

GROSSMAN, *Administrative Patent Judge*.

JUDGMENT
Final Written Decision
Determining All Challenged Claims Unpatentable
35 U.S.C. § 318(a)

I.INTRODUCTION

A. *Background and Summary*

Sonos, Inc. (“Petitioner”) filed a Petition (Paper 1, “Pet.”) requesting an *inter partes* review of claims 1–7, 9–15, 17, and 18 (the “challenged claims”) of U.S. Patent No. 10,593,330 B2 (Ex. 1001, “the ’330 patent”). Google LLC (“Patent Owner”) filed a Preliminary Response (Paper 6, “Prelim. Resp.”).

We concluded that Petitioner satisfied the burden, under 35 U.S.C. § 314(a), to show that there was a reasonable likelihood that Petitioner would prevail with respect to at least one of the challenged claims. Accordingly, on behalf of the Director (37 C.F.R. § 42.4(a) (2022)), and in accordance with *SAS Inst., Inc. v. Iancu*, 138 S. Ct. 1348, 1353 (2018), we instituted an *inter partes* review of all the challenged claims, on all the asserted grounds. Paper 9 (“Dec. Inst.”).

Patent Owner filed a Response. Paper 10 (“PO Resp.”). Petitioner filed a Reply. Paper 12 (“Reply”). Patent Owner filed a Sur-reply. Paper 13 (“Sur-reply”).

An Oral Argument was held March 5, 2024. (Paper 20) (“Transcript or “Tr.”).

We have jurisdiction under 35 U.S.C. § 6. We enter this Final Written Decision pursuant to 35 U.S.C. § 318(a) and 37 C.F.R. § 42.73.

Petitioner has the burden of proving unpatentability of a claim by a preponderance of the evidence. 35 U.S.C. § 316(e).

Based on the findings, analysis, and conclusions below, we determine that Petitioner has proven by a preponderance of the evidence that the challenged claims are unpatentable.

B. Real Parties-in-Interest

Sonos identifies itself as the sole real party-in-interest. Pet. 1. Google identifies itself as the real party-in-interest. Paper 4, 2.

C. Related Matters

The parties state that the '330 patent is the subject of the following ITC and district court proceedings: *Certain Audio Players and Components Thereof*, Inv. No. 337-TA-1329 (the "1329 ITC proceeding") and *Google LLC v. Sonos, Inc.*, Case No. 3:22-CV-04552 (N.D. Cal.) (the "District Court litigation"). Pet. 1–2; Paper 4, 2; *see also* Ex. 3004 ¶ 2 (the Complaint for Patent Infringement in the District Court litigation); Ex. 1017, 1 (Patent Owner's "Initial Markman Brief" in the 1329 ITC proceeding).

At the Oral Argument, Petitioner's counsel stated both the 1329 ITC proceeding and the District Court litigation have been stayed in view of this and related IPR proceedings. Tr. 6:16–22; *see also* Ex. 1015, 6:20–25, 13:4–13 (transcript of a Case Management Conference in the 1329 ITC proceeding discussing a revised trial schedule in light of the pending IPR proceedings).

Petitioner also states that the '330 patent "is related to the '398 Patent."¹ Pet. 2, n.1. Patent Owner acknowledges that "[t]he '330 and '398 patents are related and share a common specification." Ex. 1017, 10.

To illustrate the close relationship between the '330 and '398 patents, in addition to their common Specification, the following chart compares

¹ U.S. Patent 10,134,398 is the subject of IPR2023-00118. The '330 patent is a continuation of application No. 15/346,914, which is now the '398 patent. Ex. 1001 code (63). A Final Written Decision in IPR2023-00118 is being entered on the same day as this Decision in the case before us.

claim 1 from the '330 patent with claim 1 from the '398 patent. The italicized text in the chart identifies the differences between claim 1 in each of these patents.

CLAIM 1 – '398 PATENT IPR2023-00118	CLAIM 1 – '330 PATENT IPR2023-00119
A computer-implemented method comprising:	A computer-implemented method comprising:
receiving, by a computing device that is in a low power mode <i>and that is configured to exit a low power mode upon detecting an utterance of a particular, predefined hotword using an on-device hotword detector</i> , audio data that <i>corresponds to</i> an utterance of the particular, predefined hotword;	receiving, by a computing device that is in a low power mode, audio data that <i>includes</i> an utterance of a particular, predefined hotword;
while the computing device remains in the low power mode, and in response to receiving the audio data that <i>corresponds to</i> the utterance of the particular, predefined hotword, transmitting, by the computing device <i>and to another computing device that is configured to exit a low power mode upon detecting an utterance of the particular, predefined hotword, an output of processing the audio data using the on-device hotword detector</i> ;	while the computing device remains in the low power mode, and in response to receiving the audio data that <i>includes</i> the utterance of the particular, predefined hotword, transmitting, by the computing device, <i>a message</i> ;

while the computing device remains in low power mode, receiving, by the computing device and from <i>the other</i> computing device that is <i>configured to exit</i> a low power mode <i>upon detecting an utterance of the particular, predefined hotword</i> , an additional output of processing the audio data; and	while the computing device remains in the low power mode, receiving, by the computing device and from <i>an additional</i> computing device that is <i>in</i> a low power mode, an additional message; and
<i>after transmitting the output of processing the audio data using the on-device hotword detector and after receiving the additional output of processing the audio data from the other using device that is configured to exit a low power mode upon detecting an utterance of the particular, predefined hotword</i> , determining, by the computing device, to <i>remain</i> in the low power mode.	<i>based on the message and the additional message</i> , determining, by the computing device, to <i>exit</i> the low power mode.

Concerning related matters, Petitioner states that “Google has also asserted U.S. Patents 9,812,128 and 11,024,311 against Sonos in *Certain Audio Players and Components Thereof*, Inv. No. 337-TA-1330 and *Google LLC v. Sonos, Inc.*, Case No. 3:22-CV-04553 (N.D. Cal.)” and that “[t]hese patents are related to the ’330 Patent and have been challenged by Sonos in IPR2022-01594 and IPR2022-01592, respectively.” Pet. 2. In each of IPR2022-01592 and IPR2022-01594, the Board issued a Final Written Decision determining in each case that all challenged claims were

unpatentable. *See* IPR2022-01592, Paper 31 (PTAB Apr. 4, 2024);
IPR2022-01594, Paper 30 (PTAB Mar. 15, 2024).

D. The '330 patent

We make the following findings of fact concerning the '330 patent.

The '330 patent is titled “Hotword Detection on Multiple Devices.” Ex. 1001, code (54). The disclosed technology “generally relates to systems and techniques for recognizing the words that a person is speaking, otherwise referred to as speech recognition.” *Id.* at 1:17–19.

The disclosed technology operates in the context of “a speech-enabled home or other environment—that is, one in which a user need only speak a query or command out loud and a computer-based system will field and answer the query and/or cause the command to be performed.” *Id.* at 1:23–27.

As explained in the Specification,

[a] speech-enabled environment (e.g., home, workplace, school, etc.) can be implemented using a network of connected microphone devices distributed throughout the various rooms or areas of the environment. Through such a network of microphones, a user has the power to orally query the system from essentially anywhere in the environment without the need to have a computer or other device in front of him/her or even nearby.

Id. at 1:27–34.

A “hotword,” as used in the context of the '330 patent, is a word or phrase, which, by agreement among the users in the environment, “is reserved as a predetermined word [or phrase] that is spoken to invoke the attention of the system.” *Id.* at 1:57–61. An example of a hotword used to invoke the system’s attention is the phrase “OK computer.” Ex. 1001, 1:61–

63. In this example, each time the phrase “OK computer” is spoken within proximity of a system microphone, the system first determines whether the reserved, predetermined hotword was spoken and, “if so, awaits an ensuing command or query.” *Id.* at 1:63–67. Thus, as further explained in the ’330 patent,

utterances directed at the system take the general form [HOTWORD] [QUERY], where ‘HOTWORD’ in this example is ‘OK computer’ and ‘QUERY’ can be any question, command, declaration, or other request that can be speech recognized, parsed and acted on by the system, either alone or in conjunction with the server via the network.

Id. at 2:1–6.

According to the ’330 patent, “the system, which potentially picks up all utterances made in the surrounding environment including those not directed to the system, must have some way of discerning when any given utterance is directed at the system” and “[o]ne way to accomplish this is to use a hotword.” *Id.* at 1:52–58. Accordingly, the ’330 patent discloses an embodiment in which

a user device receives an utterance that is spoken by a user. The user device determines whether the utterance includes a hotword and computes a hotword confidence score that indicates a likelihood that the utterance includes the hotword. The user device transmits this score to other user devices in the near vicinity. The other user devices likely received the same utterance. The other user devices compute a hotword confidence score and transmit their scores to the user device. The user device compares the hotword confidence scores. If the user device has the highest hotword confidence score, then the user device *remains* active and prepares to process additional audio. If the user device does not have the highest hotword confidence score, then the user device does not process the additional audio.

Ex. 1001, 2:8–22 (emphasis added). An advantage of this system is that “[m]ultiple devices can detect a hotword and only one device will respond to the hotword.” *Id.* at 3:21–23. The system “addresses the problem of selecting the correct device for reacting to a hotword, and suppressing reaction to the hotword on other devices.” *Id.* at 3:55–57.

Figure 1 of the ’330 patent is reproduced below.

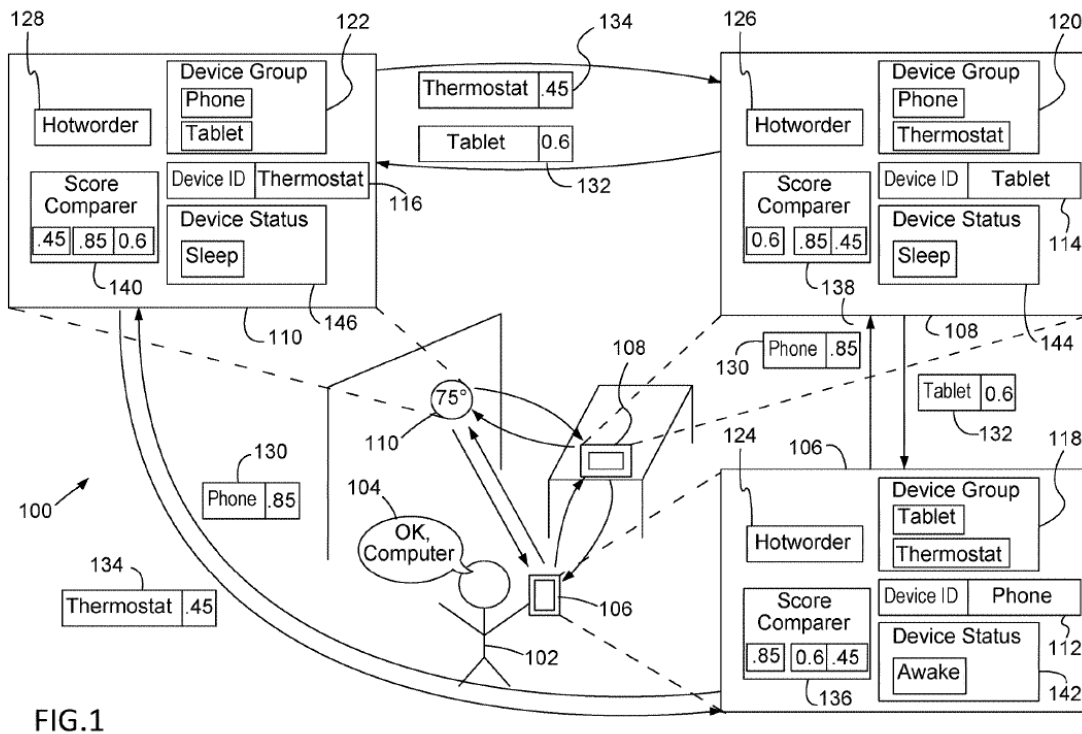


Figure 1 “is a diagram of an example system for hotword detection.” Ex. 1001, 3:33–34. System 100 allows user 102 to speak utterance 104 that is detected by microphones of computing devices 106, 108, and 110. Ex. 1001, 3:59–61. Computing devices 106, 108, and 110 process utterance 104 to determine a likelihood that utterance 104 includes a hotword and “each transmit[s] data to each other that indicates the likelihood that the utterance 104 includes a hotword.” *Id.* at 3:61–66.

Computing devices 106, 108, and 110 then compare the data and “the computing device that computed the highest likelihood that the utterance 104 included a hotword initiates speech recognition on the utterance 104” while “computing devices that did not compute the highest likelihood that the utterance 104 includes a hotword *do not initiate speech recognition* on the speech [*i.e.*, a Query] following the utterance 104.” *Id.* at 3:66–4:6 (emphasis added).

Computing devices 106, 108, and 110 include corresponding “hotworders” 124, 126, and 128 that compute a corresponding confidence score for utterance 104 from an audio input device such as a microphone using extracted audio features and “a support vector machine or a neural network.” *Id.* at 5:7–33. Computing devices 106, 108, and 110 also include corresponding score comparer 136, 138, and 140 “to compare the hotword confidence scores that the computing device has received.” *Id.* at 6:20–29. To make the comparison, confidence score data packets 130, 132, and 134 are transmitted from one computing device to another. *Id.* at 5:33–6:19.

Figure 2 from the ’330 patent, reproduced below, is a flow diagram of an example process for hotword detection. Ex. 1001, 3:35–36.

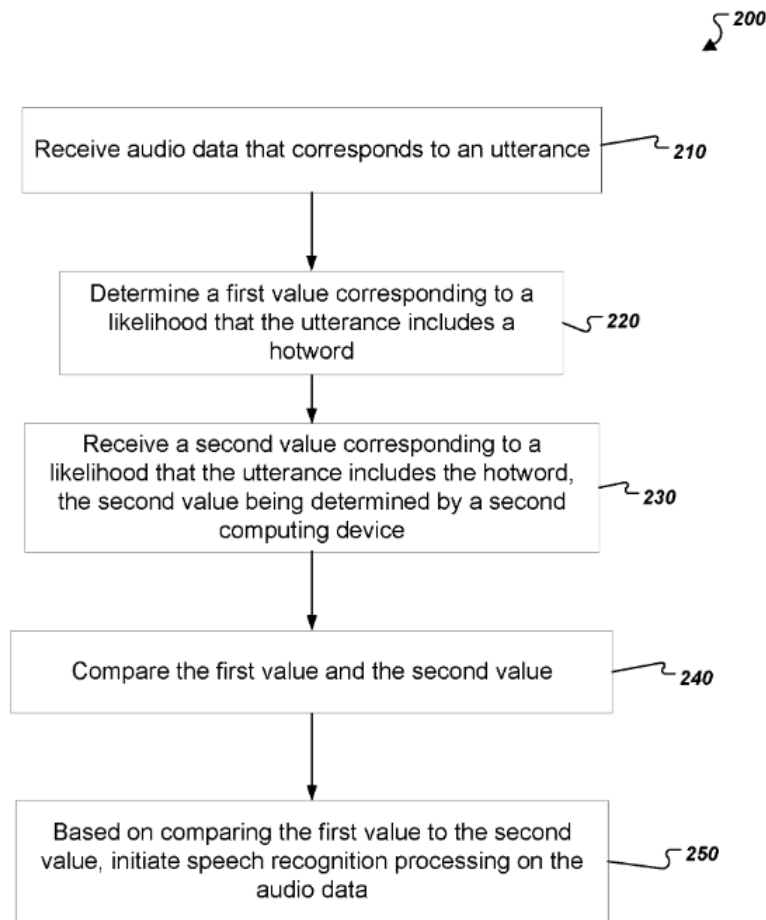


FIG. 2

Figure 2 is a diagram of an example process 200 for hotword detection that may be performed by a computing device such as computing device 108 shown in Figure 1. Ex. 1001, 9:38–41. Neither Figure 2 nor the related written description refers to a “low power mode” in determining which computing device will initiate “speech recognition” on the user query following a hotword determination.

The Specification also states that the disclosed computing devices can take many different configurations, but include typical computing elements

such as processors, memories, and similar components for networked computing devices. *See id.* at 12:24–13:12.

E. Illustrative Claim

Petitioner challenges claims 1–7, 9–15, 17, and 18. Claims 1, 9, and 17 are independent claims. Independent claim 1 is directed to a “computer-implemented method.” Ex. 1001, 16:25. Independent claim 9 is directed to a “system.” *Id.* at 17:10. Independent claim 17 is directed to a “non-transitory computer-readable medium storing software comprising instructions executable by one or more computers.” *Id.* at 18:24–26. As we discuss below in our consideration of independent claims 9 and 17, but for the preambles, claims 9 and 17 are substantive repetitions of independent claim 1.

Independent claim 1 is illustrative and is reproduced below with bracketed labels employed by Petitioner to facilitate analysis and discussion. *See* Pet. 47.

[1.0] A computer-implemented method comprising:

[1.1] receiving, by a computing device that is in a low power mode, audio data that includes an utterance of a particular, predefined hotword;

[1.2] while the computing device remains in the low power mode, and in response to receiving the audio data that includes the utterance of the particular, predefined hotword, transmitting, by the computing device, a message;

[1.3] while the computing device remains in the low power mode, receiving, by the computing device and from an additional computing device that is in a low power mode, an additional message; and

[1.4] based on the message and the additional message, determining, by the computing device, to exit the low power mode.

Ex. 1001, 16:25–40.

F. Prior Art and Asserted Grounds

Petitioner asserts that the challenged claims are unpatentable on the following grounds:

Claims Challenged	35 U.S.C. § ²	Basis
1–3, 5, 7, 9–11, 13, 15, 17, 18	102(a)	Rosenberger ³
4, 6, 12, 14	103	Rosenberger, Basye ⁴

Petitioner also relies on the declaration testimony of Dr. Michael T. Johnson. *See* Ex. 1002.⁵

Patent Owner relies on the Declaration testimony of Stuart Lipoff. *See* Ex. 2012, Ex. 2013).⁶

² The Leahy-Smith America Invents Act (“AIA”), Pub. L. No. 112-29, 125 Stat. 284, 296–07 (2011), took effect on September 16, 2011. The changes to 35 U.S.C. §§ 102 and 103 in the AIA do not apply to any patent application filed before March 16, 2013. Because the application for the patent at issue in this proceeding has an effective filing date after March 16, 2013, we refer to the AIA version of the statute.

³ U.S. Patent No. 8,340,975 B1, Dec. 25, 2012, Ex. 1003 (“Rosenberger”).

⁴ U.S. Patent Application Publ. No. 2014/01636978 A1, June 12, 2014, Ex. 1013 (“Basye”).

⁵ Dr. Johnson earned undergraduate degrees in Computer Science Engineering and in Electrical Engineering, a Master’s degree in Electrical Engineering, and a Ph.D. degree “with a doctoral dissertation focused on speech processing and automatic speech recognition.” Ex. 1002 ¶ 6. Dr. Johnson currently is the Department Chair and a Full Professor in the Department of Electrical and Computer Engineering at the University of Kentucky. *Id.* ¶ 5.

⁶ Mr. Lipoff earned undergraduate degrees in Electrical Engineering and Engineering Physics. Ex. 2012 ¶ 8. Mr. Lipoff further earned a Master of Science in Electrical Engineering and a Master of Business Administration. *Id.* Mr. Lipoff states he has been “heavily engaged in the study, analysis, evaluation, design, and implementation of products and technology

II. ANALYSIS

A. *Legal Standards*

“To anticipate a claim, a prior art reference must disclose each and every element of the claim, either explicitly or inherently.” *Adasa Inc. v. Avery Dennison Corp.*, 55 F.4th 900, 910 (Fed. Cir. 2022) (citing *Eli Lilly & Co. v. Zenith Goldline Pharms., Inc.*, 471 F.3d 1369, 1375 (Fed. Cir. 2006). “While those elements must be arranged or combined in the same way as in the claim, the reference need not disclose the elements in the very same terms used by the patent.” *Id.* (citing *In re Gleave*, 560 F.3d 1331, 1334 (Fed. Cir. 2009) (“[T]he reference need not satisfy an *ipsissimis verbis* test.” (citing *In re Bond*, 910 F.2d 831, 832–33 (Fed. Cir. 1990))). “The question of what a reference teaches and whether it describes every element of a claim is a question for the finder of fact.” *Id.* (citing *Med. Instrumentation & Diagnostics Corp. v. Elekta AB*, 344 F.3d 1205, 1221 (Fed. Cir. 2003)). As further explained in *Net MoneyIN*,

unless a reference discloses within the four corners of the document not only all of the limitations claimed but also all of the limitations arranged or combined in the same way as recited in the claim, it cannot be said to prove prior invention of the thing claimed and, thus, cannot anticipate under 35 U.S.C. § 102.

Net MoneyIN, Inc. v. VeriSign, Inc., 545 F.3d 1359, 1371 (Fed. Cir. 2008). *Id.*; see also *Verdegaal Bros. v. Union Oil Co.*, 814 F.2d 628, 631 (Fed. Cir. 1987) (“A claim is anticipated only if each and every element as set forth in

associated with consumer electronics and electronic appliances.” *Id.* ¶ 16. A particular focus of his professional activities has been “improving the man-machine interface including voice, speech, and speaker recognition for man-machine interactions.” *Id.*

the claim is found, either expressly or inherently described, in a single prior art reference.”). “The identical invention must be shown in as complete detail as is contained in the . . . claim.” *Richardson v. Suzuki Motor Co.*, 868 F.2d 1226, 1236 (Fed. Cir. 1989).

Whether the challenged claims would have been obvious over the cited references also is at issue in this proceeding. Section 103 forbids issuance of a patent when “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.” *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations, including: (1) the scope and content of the prior art; (2) any differences between the claimed subject matter and the prior art; (3) the level of ordinary skill in the art; and (4) when available, evidence such as commercial success, long felt but unsolved needs, and failure of others.⁷ *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966); *see KSR*, 550 U.S. at 407 (“While the sequence of these questions might be reordered in any particular case, the [Graham] factors continue to define the inquiry that controls.”). The Court in *Graham* explained that these factual inquiries promote “uniformity and definiteness,” for “[w]hat is obvious is not a question upon which there is likely to be uniformity of thought in every given factual context.” 383 U.S. at 18.

⁷ Patent Owner does not direct us to any persuasive objective evidence of non-obviousness.

The Supreme Court made clear that we apply “an expansive and flexible approach” to the question of obviousness. *KSR*, 550 U.S. at 415. Whether a patent claiming the combination of prior art elements would have been obvious is determined by whether the improvement is more than the predictable use of prior art elements according to their established functions. *Id.* at 417. To support this conclusion, however, it is not enough to show merely that the prior art includes separate references covering each separate limitation in a challenged claim. *Unigene Labs., Inc. v. Apotex, Inc.*, 655 F.3d 1352, 1360 (Fed. Cir. 2011). Rather, obviousness additionally requires that a person of ordinary skill at the time of the invention “would have selected and combined those prior art elements in the normal course of research and development to yield the claimed invention.” *Id.*

In determining whether there would have been a motivation to combine prior art references to arrive at the claimed invention, it is insufficient to simply conclude the combination would have been obvious without identifying any reason *why* a person of skill in the art would have made the combination. *Metalcraft of Mayville, Inc. v. Toro Co.*, 848 F.3d 1358, 1366 (Fed. Cir. 2017).

Moreover, in determining the differences between the prior art and the claims, the question under 35 U.S.C. § 103 is not whether the differences themselves would have been obvious, but whether the claimed invention as a whole would have been obvious. *Litton Indus. Prods., Inc. v. Solid State Sys. Corp.*, 755 F.2d 158, 164 (Fed. Cir. 1985) (“It is elementary that the claimed invention must be considered as a whole in deciding the question of obviousness.”); *see also Stratoflex, Inc. v. Aeroquip Corp.*, 713 F.2d 1530, 1537 (Fed. Cir. 1983) (“[T]he question under 35 U.S.C. § 103 is not whether

the differences *themselves* would have been obvious. Consideration of differences, like each of the findings set forth in *Graham*, is but an aid in reaching the ultimate determination of whether the claimed invention as a whole would have been obvious.”).

As a factfinder, we also must be aware “of the distortion caused by hindsight bias and must be cautious of arguments reliant upon *ex post* reasoning.” *KSR*, 550 U.S. at 421.

Applying these general principles, we consider the evidence and arguments of the parties.

B. Level of Ordinary Skill in the Art

The level of skill in the art is “a prism or lens” through which we view the prior art and the claimed invention. *Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001). “This reference point prevents . . . factfinders from using their own insight or, worse yet, hindsight, to gauge obviousness.” *Id.*

“The *Graham* analysis includes a factual determination of the level of ordinary skill in the art. Without that information, a . . . court cannot properly assess obviousness because the critical question is whether a claimed invention would have been obvious at the time it was made to one with ordinary skill in the art.” *Custom Accessories, Inc. v. Jeffrey-Allan Indus., Inc.*, 807 F.2d 955, 962 (Fed. Cir. 1986); *see also Ruiz v. A.B. Chance*, 234 F.3d 654, 666 (Fed. Cir. 2000) (“The determination of the level of skill in the art is an integral part of the *Graham* analysis.”).

Factors pertinent to a determination of the level of ordinary skill in the art include: (1) educational level of the inventor; (2) type of problems encountered in the art; (3) prior art solutions to those problems; (4) rapidity

with which innovations are made; (5) sophistication of the technology; and (6) educational level of workers active in the field. *Best Med. Int'l, Inc. v. Elekta Inc.*, 46 F.4th 1346, 1353 (Fed. Cir. 2022) (citing *Daiichi Sankyo Co. v. Apotex, Inc.*, 501 F.3d 1254, 1256 (Fed. Cir. 2007) (quoting *Env't Designs, Ltd. v. Union Oil Co.*, 713 F.2d 693, 696 (Fed. Cir. 1983)). These factors are not exhaustive but merely are a guide to determining the level of ordinary skill in the art. *Daiichi Sankyo*, 501 F.3d at 1256. In determining a level of ordinary skill, we also may look to the prior art, which may reflect an appropriate skill level. *Okajima*, 261 F.3d at 1355.

Petitioner asserts that a person of ordinary skill in the art would have had “the equivalent of a four-year degree from an accredited institution (typically denoted as a B.S. degree) in electrical or computer engineering, or a related field, and approximately 2–4 years of experience in the field of voice and/or audio processing, or an equivalent level of skill, knowledge, and experience.” Pet. 22 (citing Ex. 1002 ¶¶ 59–62).

Dr. Johnson’s Declaration testimony states he considered some of the factors listed above. Ex. 1002 ¶ 60. Dr. Johnson “also considered [his] own experiences working and teaching in the fields of voice and audio processing.” *Id.* Dr. Johnson testifies that, in his opinion, the level of ordinary skill is as stated by Petitioner. *Id.* at 61.

Patent Owner does not assert a level of ordinary skill, nor does Patent Owner dispute or otherwise comment on Petitioner’s proposed definition of the level of ordinary skill. *See* PO Resp. and Sur-reply, *generally*.

Based on the prior art, the sophistication of the disclosed technology in the ’330 patent, and Dr. Johnson’s testimony, we adopt Petitioner’s definition of the level of ordinary skill. Accordingly, based on the evidence

before us, we determine a person of ordinary skill in the context of the '330 patent would have had the equivalent of a four-year degree from an accredited institution (typically denoted as a B.S. degree) in electrical or computer engineering, or a related field, and approximately 2–4 years of experience in the field of voice and/or audio processing, or an equivalent level of skill, knowledge, and experience.

C. Claim Construction

We construe each claim “using the same claim construction standard that would be used to construe the claim in a civil action under 35 U.S.C. [§] 282(b).” 37 C.F.R. § 42.100(b). Under this standard, claim terms are generally given their ordinary and customary meaning as would have been understood by a person of ordinary skill in the art at the time of the invention and in the context of the entire patent disclosure. *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312–14 (Fed. Cir. 2005) (en banc) (“We have frequently stated that the words of a claim ‘are generally given their ordinary and customary meaning.’” (citations omitted)); *see also Grace Instrument Indus., LLC v. Chandler Instruments Co., LLC*, 57 F.4th 1001, 1008 (Fed. Cir. 2023) (quoting *Phillips*, 415 F.3d at 1313). While there is substantial judicial precedent and guidance concerning claim construction, “there is no magic formula or catechism for conducting claim construction.” *Intel Corp. v. Qualcomm Inc.*, 21 F.4th 801, 809 (Fed. Cir. 2021) (quoting *Phillips*, 415 F.3d at 1324).

The Specification, or more precisely, the written description, is the “single best guide to the meaning of a *disputed* term.” *Grace Instrument*, 57 F.4th at 1008 (quoting *Vitronics Corp. v. Conceptoronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996) (emphasis added), and “is, thus, the primary basis for

construing the claims.” *Id.* (citation omitted). This focus on the Specification helps to avoid what has been called “the curse of indefinite and ambiguous claims, divorced from the written description.” *Retractable Techs., Inc. v. Becton, Dickinson & Co.*, 653 F.3d 1296, 1311 (Fed. Cir. 2011) (Plager, Circuit Judge, concurring).

Petitioner states, “the Challenged Claims do not require construction.” Pet. 22.

In its Preliminary Response, Patent Owner did not propose any claim construction, nor did Patent Owner comment on Petitioner’s statement that the challenged claims do not require construction. *See* Prelim. Resp., *generally*.

Accordingly, in our Institution Decision, we determined that, “[w]e do not perceive the need to construe any claim limitation.” Dec. Inst. 14.

In its Response, however, Patent Owner proposes a specific construction for the term “low power mode.” PO Resp. 2–6. The term “low power mode” is used in all three independent claims (claims 1, 9, and 17). This term also is used specifically in challenged dependent claims 3, 7, 11, and 15.

According to Patent Owner, the term “low power mode” “has a well-understood plain and ordinary meaning: ‘an operating mode or state in which power is conserved.’” PO Resp. 2–3. Patent Owner states “[t]his construction is consistent with (1) the patent’s prosecution history and the

cited art, (2) the patent’s specification⁸, and (3) the extrinsic evidence.” PO Resp. 3 (citing Ex. 2012 ¶¶ 54–58).⁹

In its Reply, Petitioner notes that “Google concedes that Rosenberger discloses [a] *low power mode* under its interpretation.” Reply 1 (citing PO Resp. 2 (“Google also does not dispute that Rosenberger’s ‘low-power ‘listening mode’ is a low power mode, as claimed in the ’330 Patent....”). Petitioner, therefore, states that “the Board can assume that Google’s proposed interpretation of low power mode is correct, which means there is no controversy for the Board to resolve.” *Id.*; *see also* Sur-reply 1 (“Petitioner does not dispute Patent Owner’s construction of “low power mode”).

⁸ The claims of the ’330 patent use the term “low power mode” twenty times. This term, however, does *not* appear in the written description of the ’330 patent. In its Sur-reply, Patent Owner states “the ’330 patent expressly discloses ‘active’ and ‘inactive’ states—the latter of which clearly corresponds to an operating mode or state in which power is conserved.” Sur-reply 2 n.1. Patent Owner cites no persuasive evidence to support this argument. “Argument in the brief does not take the place of evidence in the record.” *Garrido v. Holt*, 547 F. App’x 974, 979 n.3 (Fed. Cir. 2013), (citing *In re Schulze*, 346 F.2d 600, 602 (CCPA 1965).

We also note that in the related 1329 ITC proceeding, the Judge stated she would be issuing a “Markman order” for the term “low power mode” (Ex. 1015, 6:9–13) with a determination that this term is “indefinite” (*id.* at 9:17–20). The issue of whether the term “low power mode” is “indefinite” under 35 U.S.C. § 112 is beyond our jurisdiction in this IPR proceeding, which is limited to determinations under “[35 U.S.C. §§] 102 or 103 and only on the basis of prior art consisting of patents or printed publications.” *See* 35 U.S.C. § 311.

⁹ Patent Owner cites the Declaration testimony of Mr. Lipoff as “Decl.” We refer to it, as we do with all evidence before us, by its exhibit number, which is Ex. 2012.

Claims must be construed “only to the extent necessary to resolve the controversy.” *AstraZeneca AB v. Mylan Pharms. Inc.*, 19 F.4th 1325, 1345 (Fed. Cir. 2021) (quoting *Vivid Techs., Inc. v. American Science & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999); *see also Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (“we need only construe terms ‘that are in controversy, and only to the extent necessary to resolve the controversy’” (citations omitted)). We determine that a specific construction of the term “low power mode” is not necessary to reach a decision in this *inter partes* review of the challenged claims.

Both parties agree that the ordinary and customary meaning of “low power mode” is “an operating mode or state in which power is conserved,” and thus, this is the construction we will use in reaching our patentability determination. As we discuss below, the dispute between the parties concerns *when* a networked device in Rosenberger *exits* the low power mode, *not whether* Rosenberger discloses a low power mode as claimed. As stated by Patent Owner, the dispositive issue in this proceeding is whether Rosenberger “teach[es] a process by which devices in a low power mode remain in that low power mode while exchanging messages with other devices for the purpose of determining whether to exit the low power mode.” PO Resp. 2.

We now turn to the merits of Petitioner’s asserted Grounds of unpatentability.

*D. Ground 1 – Anticipation of Claims
1–3, 5, 7, 9–11, 13, 15, 17, and 18
Based on Rosenberger*

Petitioner asserts claims 1–3, 5, 7, 9–11, 13, 15, 17, and 18 are anticipated by Rosenberger. Pet. 22–60.

First, we summarize the disclosure of Rosenberger.

1. Rosenberger (Ex. 1003)

We make the following findings regarding Rosenberger.

Rosenberger is titled “Interactive Speech Recognition Device and System for Hands-Free Building Control.” Ex. 1003, code (54). Rosenberger “relates to improvements in automated control systems” and, “[m]ore particularly, . . . to improvements in speech-recognizing control devices that respond to oral commands from a system user.” *Id.* at 1:8–13.

As summarized in its Abstract, Rosenberger discloses:

A self-contained wireless interactive speech recognition control device and system that integrates with automated systems and appliances to provide totally hands-free speech control capabilities for a given space. Preferably, each device comprises a programmable microcontroller having embedded speech recognition and audio output capabilities, a microphone, a speaker and a wireless communication system through which a plurality of devices can communicate with each other and with one or more system controllers or automated mechanisms. The device may be enclosed in a standalone housing or within a standard electrical wall box. Several devices may be installed in close proximity to one another to ensure hands-free coverage throughout the space. When two or more devices are triggered simultaneously by the same speech command, real time coordination ensures that only one device will respond to the command.

Ex. 1003, code (57).

As disclosed in Rosenberg, to provide complete hands-free speech recognition coverage and comfortable user interaction from anywhere in a building using speech recognition modules of this type, it would often be necessary to install many modules in reasonable proximity to one another. *Id.* at 2:42–47. In such an arrangement, it is inevitable that a given speech command would frequently be recognized simultaneously by two or more neighboring speech recognition modules. *Id.* at 2:47–50. If multiple modules were triggered and responded simultaneously, the target automated mechanism could be instructed by each triggered device to perform the same instruction. *Id.* at 2:50–56. For example, the speech command “Lower Room Temperature Two Degrees” could be executed separately by several modules causing a thermostat to lower by multiples of the requested two degrees. *Id.* at 2:56–59. This is referred to as a “duplicate response problem.” *Id.* at 2:60. One of the objectives of the system disclosed in Rosenberg is to avoid the “duplicate response problem.” *Id.* at 3:30–35.

Rosenberger discloses “a totally hands-free, interactive, wireless speech recognition device and system that operate using low-power, low-data rate, wireless connections.” *Id.* at 3:45–49. The disclosed system allows the user to wirelessly control the operation of automated mechanisms using speech commands, so that “[w]hen a trigger phrase is detected, the system ‘wakes up’ and prompts the user to say one or more commands from an expected command phrase vocabulary.” *Id.* at 1:63–2:9.

Figure 3 of Rosenberg is reproduced below.

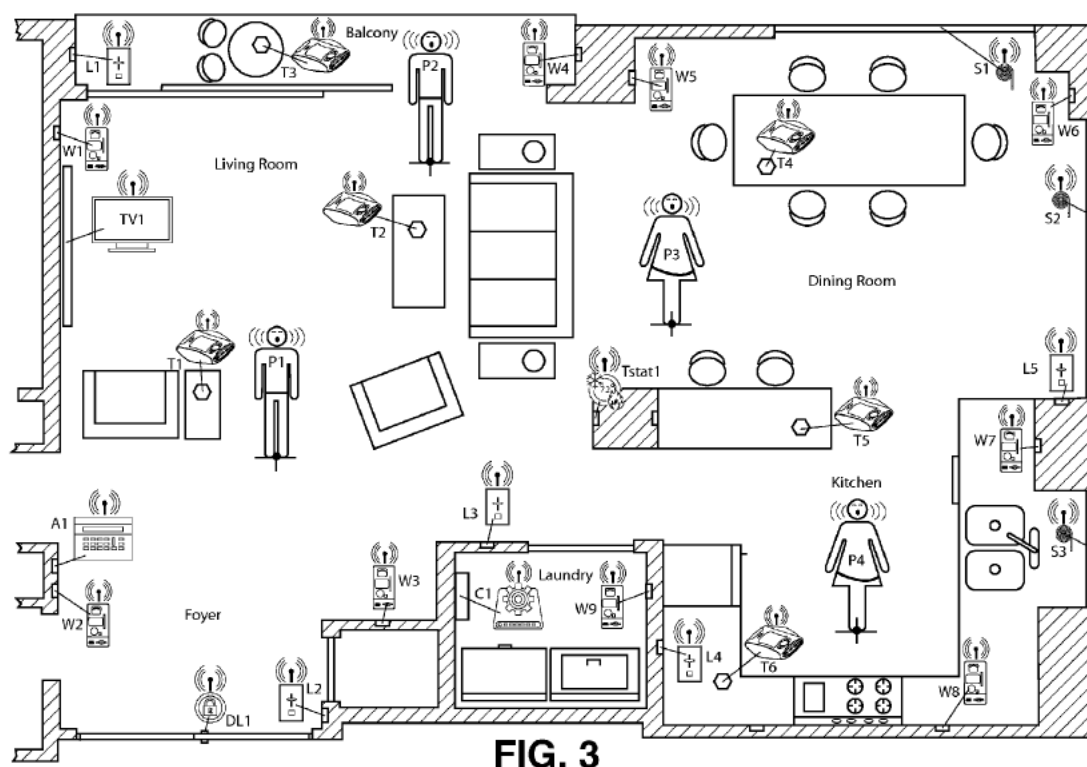


Figure 3 “illustrates a residential floor plan showing a plurality of interactive speech recognition control devices at various locations that provide totally hands-free speech control of various automated mechanisms throughout the interior space.” Ex. 1003, 5:41–44. A home-automation system includes table-top control devices T1–T6, electrical wall-box mounted control devices W1–W9, and automated mechanisms such as door lock DL1, security panel A1, television TV1, thermostat Tstat1, light dimmers L1–L5, and motorized shades S1–S3, that may all be controlled “hands-free” by persons P1–P4 using only speech commands. *Id.* at 9:37–49. “Control signals may be sent directly from any speech recognition device to system controller C1, which in turn, sends appropriate control signals to the target automated mechanisms to implement the speech command” or “any speech recognition device can control the automated mechanisms directly using

standard wireless communications and automation protocols.” *Id.* at 9:50–56. Any of the table top devices T1–T6 and wall-boxes mounted control devices W1–W9 “may ‘hear’ and recognize” a “speech trigger” to operate the automated mechanisms. *Id.* at 9:57–10:4.

Rosenberger discloses that the control device “uses one of several different coordination methods to achieve the design goal of avoiding duplicate responses to the same speech command while selecting a single control device best able to successfully interact with the system user” and “[a]ll of the methods involve the calculation and production by the control device of a numeric weighted signal (WS) that quantifies the ‘quality’ of a just-recognized speech command.” *Id.* at 10:34–41.

As disclosed in Rosenberger,

Preferably, device D [a control device] *normally operates in a low power ‘listening mode’ in which it listens for a hands-free trigger phrase* (e.g., “Hello Voice Control or “Home Automation”) to be spoken by the system user. *Upon receiving and recognizing a speech trigger phrase [Rosenberger’s label for a “Hotword”] and determining that it is in a better position to handle subsequent user interaction than any other device that simultaneously recognized the same speech trigger phrase* (see device coordination discussion below), or when the user depresses the “Push to Talk” button 24 on the device, or in the event the device is *instructed to wake up and expect a subsequent speech command by a system controller*, the device beeps, *changes its status light 34 [see Figs. 1, 2B]* or plays a prerecorded or synthesized audio message (e.g. “How may I help you?”) through speaker 19 and/or 20 to prompt the user to say one of the speech commands from a known vocabulary or grammar stored in the speech pattern memory 14. Upon recognizing a particular speech command (and optionally confirming with the user), the device produces and transmits either to a system controller or directly to the automated mechanisms (via its communications circuitry 18) a corresponding system control signal designed to

instruct the appropriate automated mechanisms to carry out the subject of the speech command.

Ex. 1003, 8:17–39 (emphases added).

As disclosed, Rosenberger’s “low power listening mode” is for listening for a trigger phrase. *Id.* at 8:17–19. It is only *after* detecting the trigger phrase and “*upon . . . determining that it is in a better position to handle subsequent user interaction than any other device that simultaneously recognized the same speech trigger phrase*” does Rosenberger’s device D “*change[] its status light 34*” to indicate its new “wake up” status as being ready for a subsequent speech command. *Id.* at 8:20–27 (emphases added).

Thus, to summarize the disclosures cited above, Rosenberger’s computing devices (1) listen in “low power” mode for a “trigger phrase” or “Hotword;” (2) upon detecting a trigger phrase, the devices exchange information in the form of “weighted signals” to determine which computing device is “in a better position” to handle a subsequent user query; and (3) that one device in a “better position” than the other devices then exits its low power mode and “wakes up” so it can respond to any subsequent user query. How does the computing device accomplish these numerous and diverse tasks? Figure 1 from Rosenberger, showing the components within each of Rosenberger’s computing devices D, provides an illustrated answer, and is reproduced below.

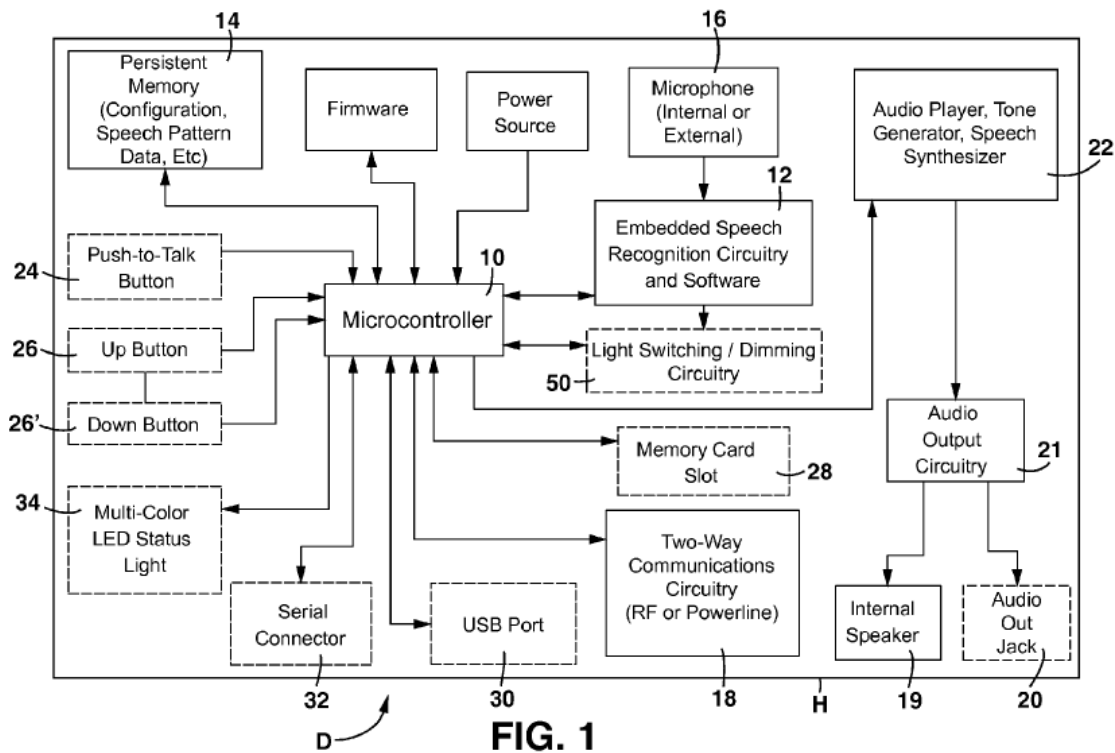


Figure 1 from Rosenberger “schematically illustrates an interactive speech recognition control device D that is configured in accordance with a preferred embodiment” of Rosenberger’s disclosed invention. Ex. 1003, 6:40–43; *see also id.* Fig. 2A (Fig.. 2A “is a perspective illustration of a contemplated table-top housing for containing the device of FIG. 1”).

In order to achieve the described performance of its computing devices, Rosenberger discloses that the computing devices include “a uniquely-programmed microcontroller 10 having embedded speech recognition circuitry and software 12. *Id.* at 6:62–65. Each computing device also includes:

a persistent internal memory 14 (e.g. non-volatile flash memory) for storing speech pattern data (phoneme expressions, etc.) and the like, a microphone 16 through which speech commands from the user are received, and a two-way, preferably wireless, communication system 18, all being supported by a housing H

that can be placed anywhere in a room or installed in an electrical wall box.

Ex. 1003, 6:65–7:4.

As further explained in Rosenberger,

each control device D is *programmed* so that multiple devices can be placed in close proximity to one another in the same space. In accordance with a preferred embodiment, different devices recognizing the same speech command from the same person at the same instant *coordinate with each other* (shown in FIGS. 9, 10), or with an external coordinating controller CC (shown in FIGS. 7, 8), to ensure that only one device handles each discrete speech command, preferably the one able to hear the user as loudly and clearly as possible.

Id. at 10:5–14 (emphases added).

Rosenberger’s control devices use one of several different coordination methods to achieve the design goal of avoiding duplicate responses to the same speech command while selecting a single control device best able to successfully interact with the system user. Ex. 1003, 10:34–38. All of Rosenberger’s disclosed methods involve “the calculation and production by the control device of a numeric weighted signal (WS) that quantifies the ‘quality’ of a just-recognized speech command.” *Id.* at 10:38–41.

2. *Independent Claims 1, 9, 17*

Petitioner argues claims 1, 9, and 17 collectively. *See* Pet. 22–49. Petitioner states that the “Petition addresses independent claim 9 first because it is the narrowest independent claim of the ’330 Patent.” *Id.* at 22. We make no finding or conclusion about the scope of claim 9 compared to claims 1 and 17. For consistency with the Petition, however, we too will

begin our analysis with independent claim 9. Before doing so, we provide the chart below showing the substantial similarity of claims 1, 9, and 17.

In the chart below, we use the paragraph numbering for these claims used by Petitioner. As is evident, the only significant difference among these three claims is their preamble.

CLAIM 1	CLAIM 9	CLAIM 17
[1.0] A computer-implemented method comprising:	[9.0] A system comprising: one or more computers; and one or more computers and one or more storage devices storing instructions that are operable, when executed by the one or more computers, to cause the one or more computers to perform operations comprising:	[17.0] A non-transitory computer-readable medium storing software comprising instructions executable by one or more computers which, upon such execution, cause the one or more computers to perform operations comprising:
[1.1] receiving, by a computing device that is in a low power mode, audio data that includes an utterance of a particular, predefined hotword;	[9.1] receiving, by a computing device that is in a low power mode, audio data that includes an utterance of a particular, predefined hotword;	[17.1] receiving, by a computing device that is in a low power mode, audio data that includes an utterance of a particular, predefined hotword;

[1.2] while the computing device remains in the low power mode, and in response to receiving the audio data that includes the utterance of the particular, predefined hotword, transmitting, by the computing device, a message;	[9.2] while the computing device remains in the low power mode, and in response to receiving the audio data that includes the utterance of the particular, predefined hotword, transmitting, by the computing device, a message;	[17.2] while the computing device remains in the low power mode, and in response to receiving the audio data that includes the utterance of the particular, predefined hotword, transmitting, by the computing device, a message;
[1.3] while the computing device remains in the low power mode, receiving, by the computing device and from an additional computing device that is in a low power mode, an additional message; and	[9.3] while the computing device remains in the low power mode, receiving, by the computing device and from an additional computing device that is in a low power mode, an additional message; and	[17.3] while the computing device remains in the low power mode, receiving, by the computing device and from an additional computing device that is in a low power mode, an additional message; and
[1.4] based on the message and the additional message, determining, by the computing device, to exit the low power mode.	[9.4] based on the message and the additional message, determining, by the computing device, to exit the low power mode.	[17.4] based on the message and the additional message, determining, by the computing device, to exit the low power mode.

Petitioner provides a clause-by-clause analysis of claim 9 mapping each element or limitation of claim 9 to what Petitioner asserts is the corresponding disclosure in Rosenberger. Pet. 22–46.

Throughout its analysis of claims 1, 9, and 17, Petitioner cites to the written description and drawings of Rosenberger and to the Declaration testimony of Dr. Johnson for evidentiary support. *See, e.g., id.* at 24 (citing Ex. 1002 ¶ 74), 26 (citing Ex. 1002 ¶ 77).

In general summary, and focusing on the issues argued by the parties, independent claims 1, 9, and 17 each require three actions while the claimed “computing devices” are in a low power mode.

First, “a computing device” (a first computing device) in a low power mode *receives* audio data that “includes an utterance” of a predefined hotword. Ex. 1001, 16:26–28.¹⁰ This step, element, or limitation is contained in clauses 1.1, 9.1, and 17.1.

Second, while the first computing device remains in the low power mode, and in response to receiving the audio data, it *transmits* a “message.” *Id.* at 16:29–33. This step, element, or limitation is contained in clauses 1.2, 9.2, and 16.2. As disclosed in the Specification, one form of this claimed “message” may be “a hotword confidence score that indicates a likelihood that the utterance includes the hotword.” *Id.* at 2:6–12.

Third, while the first computing device remains in the low power mode, it *receives* from an “additional computing device” (a second computing device) an “additional message.” Ex. 1001, 16:34–37. This step, element, or limitation is contained in clauses 1.3, 9.3, and 17.3. This “additional message” from the second computing device also may be “a hotword confidence score that indicates a likelihood that the utterance includes the hotword.” *Id.* at 2:6–12.

¹⁰ We cite to claim 1. As shown in the chart comparing the claims, these same steps, elements, or limitations are recited in claims 9 and 17.

At this point in the claimed invention, the computing devices, still in low power mode, determine, based on the “message” and the “additional message,” whether to “exit the low power mode.” Ex. 1001, 16:38–40. We know from the Specification that this determination is made by comparing “hotword confidence scores”, with the computing device with the higher score exiting the low power mode and going into an active mode ready to receive and respond to a user query, while the other computing device remains in low power mode. Ex. 1001, 2:6–21. We do *not* read this disclosure from the Specification into the claims. We mention it to provide context for what the claims recite, and what they do *not* recite.

We now turn to specific clauses and arguments regarding the independent claims, beginning, as did Petitioner, with claim 9.

a. Claim 9

i. Clause [9.0]

[9.0] A system comprising: one or more computers;¹¹

Petitioner relies on Figures 1 and 3 (reproduced and discussed above) and the related disclosure in Rosenberger to establish that Rosenberger discloses the system in clause [9.0]. Pet. 23–27 (citing Ex. 1002 ¶¶ 72–83).

Petitioner asserts a person of ordinary skill in the relevant technology, which we have determined is a person with the equivalent of a four-year degree from an accredited institution in electrical or computer engineering, or a related field, and approximately 2-4 years of experience in the field of

¹¹ Petitioner clearly misstates clause [9.0] in the Petition at pages 22–23 in its clause-by-clause analysis of claim 9. We have ignored this clear error and rely on the correct preamble language in the ’330 patent. Dr. Johnson makes this same error in his Declaration testimony. Ex. 1002 ¶¶ 71, 72.

voice and/or audio processing (*see* Section II.B of this Decision), “would have understood that each ‘control device’ in Rosenberger amounts to *a computer that comprises one or more storage devices storing instructions that are operable, when executed by the [computer], to cause the [computer] to perform operations.*” Pet. 24 (citing Ex. 1002 ¶ 74 (Dr. Johnson’s opinion testimony supporting Petitioner’s assertion); *id.* at 26 (citing Ex. 1002 ¶ 82); Ex. 1014, 15–18 (Patent Owner’s infringement contention explaining why the accused products are “computers,” as claimed in the ’330 patent)).

The claims in Rosenberger also are part of Rosenberger’s disclosure and Specification. 35 U.S.C. § 112(b) (“The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the inventor or a joint inventor regards as the invention.”). Claim 1 in Rosenberger, for example, provides a summary of the elements of each of Rosenberger’s “functionally identical [control] devices” (Ex. 1003, 17:12–13). These elements and their functions include:

- (a) a microphone for converting spoken words, emanating from a system user located within said building space, to corresponding electrical signals;

- (b) a programmable microcontroller, operatively coupled to the microphone and comprising embedded software and circuitry . . . ;

- (c) a transmitter operatively coupled to said microcontroller for selectively transmitting said weighted signal and said system control signals into said building space;

- (d) a receiver, operatively coupled to said microcontroller, for receiving (i) weighted signals emanating from one or more other functionally identical interactive speech recognition control devices within said building space, and/or (ii)

coordination signal derived from said weighted signals by an independent coordinating controller; and

(e) audio output circuitry through which said microcontroller can transmit audible sounds to interact with and provide indications of the system status to the system user.

Ex. 1003, 17:17–52; *see also* Pet. 28–29 (citing Ex. 1003, 17:10–27).

It is beyond reasonable dispute that Rosenberger’s “control devices D” are computers as stated in claim 9 of the ’330 patent. Indeed, Patent Owner does not take specific issue with the preamble clause, clause 9.0, and the corresponding clauses in independent claims 1 and 17.

Accordingly, based on a preponderance of the evidence before us, we determine that Rosenberger discloses the steps, elements and limitations of clause [9.0].

ii. Clause [9.1]

[9.1] receiving, by a computing device that is in a low power mode, audio data that includes an utterance of a particular, predefined hotword;

Petitioner asserts that Rosenberger discloses the steps, elements, and limitations in clause [9.1]. Pet. 27–31. Throughout its analysis of this clause, Petitioner cites the Declaration testimony of Dr. Johnson for evidentiary support. *See, e.g., id.* (citing Ex. 1002 ¶¶ 84–97). Petitioner also cites Rosenberger’s claim 1 for further evidence of how Rosenberger operates. Pet. 29 (citing Ex. 1003, 17:10–27).

Rosenberger’s trigger phrase is a “hotword.” Dr. Johnson testifies that:

Rosenberger explains that, in preferred embodiments, each ‘control device’ operates in a “low power ‘listening’ mode in which it listens for a hands-free trigger phrase (e.g., ‘Hello Voice Control’ or ‘Home Automation’) to be spoken by the system

user’ and *then only one* given “control device” exits that ‘low power’ mode to interact with the user and process voice commands after employing one of the disclosed ‘coordination methods’ and determining that the given ‘control device’ is in ‘a better position to handle subsequent user interaction than any other device that simultaneously recognized the same speech trigger phrase’

Ex. 1002 ¶ 47 (emphasis added) (citing Ex. 1003, 8:15–32).

Patent Owner’s basic argument as to why Rosenberger does not anticipate independent claims 1, 9, and 17 is that “Rosenberger does not teach a process by which devices in a low power mode *remain in that low power mode* while exchanging messages with other devices for the purpose of determining whether to remain in the low power mode.” PO Resp. 2 (emphasis added). We discuss this in the following sections.

Patent Owner’s basic argument, however, assumes at least two or more “computing devices” that exchange messages with other computing devices for the purpose of determining which device will exit the lower power mode and respond to a user query, while the other device or devices remain in the low power mode. PO Resp. 2. If there is only one computing device in the system, there is no exchange of data or exchange of a “message” with other computing devices in the system to determine which device will exit the low power mode. Thus, Patent Owner’s basic argument does not apply to clause [9.1] because the claimed computing device is not yet exchanging and/or comparing data with any other computing device.

Accordingly, based on a preponderance of the evidence before us, we determine that Rosenberger discloses the steps, elements and limitations of clause [9.1].

iii. Clause [9.2]

[9.2] while the computing device remains in the low power mode, and in response to receiving the audio data that includes the utterance of the particular, predefined hotword, transmitting, by the computing device, a message;

Clause [9.2] requires that “while the computing device remains in the low power mode,” it transmits “a message.” This clause does not state to what or whom it transmits this message, nor does it state the type or content of the message.

Petitioner asserts that Rosenberger discloses the steps, elements, and limitations in clause [9.2]. Pet. 31–39 (citing Ex. 1002 ¶¶ 98–115).

Petitioner asserts that “Rosenberger discloses that each ‘control device’ is configured to¹² calculate and produce a ‘weighted signal’ (WS) value in response to receiving audio data corresponding to a user’s utterance containing a ‘trigger phrase.’” Pet. 31–34. According to Petitioner,

Rosenberger explains that each ‘control device’ ‘sends its weighted signal to each of the other interactive speech recognition control devices connected to the same network using a broadcast or group multicast **messaging** technique’ and ‘[i]mmediately after broadcasting its weighted signal, a triggered [control] device waits a predetermined amount of time while it monitors **weighted signal messages** that may have been simultaneously broadcasted from other like devices.

¹² Patent Owner argues that Petitioner argues that Rosenberger is “capable of” performing certain functions, which, according to Patent Owner “is insufficient to establish anticipation.” PO Resp. 10–13. While the term “capable of” may appear in some of Petitioner’s assertions, it is clear that both Petitioner and Rosenberger discuss and disclose, respectively, a system and method that operates and performs the functions disclosed.

Pet. 34 (citing Ex. 1003, 11:45–59 (emphasis in Petition)). Thus, it is Petitioner’s position that this weighted signal is the “message” include in the claims.

Patent Owner argues that Rosenberger does not expressly disclose this limitation. PO Resp. 10. Patent Owner’s basic argument as to why Rosenberger does not anticipate independent claims 1, 9, and 17 is that “Rosenberger does not teach a process by which devices in a low power mode *remain in that low power mode* while exchanging messages with other devices for the purpose of determining whether to remain in the low power mode.” PO Resp. 2 (emphasis added). It is Patent Owner’s position that “Rosenberger expressly discloses that its devices exit the low power ‘listening mode’ when they are ‘triggered by a speech command,’ *prior to exchanging messages* and selecting which device will respond to the user.” *Id.* (emphasis added). As further explained by Patent Owner,

each independent claim [in the ’330 patent] recites a process by which a device that receives audio data including an utterance of a particular, predefined hotword while in a low power mode, remains in that low power mode while exchanging outputs with another device operating in a low power mode. Upon receipt of an additional output from another device, the device determines whether to remain in the low power mode.

Id. at 6.

Patent Owner asserts

Rosenberger does not anticipate claim 1 because Rosenberger discloses sending a “weighted signal” (the alleged “message”) after the control device (the alleged “computing device”) is “triggered,” “i.e., switched from a ‘listening’ mode” (the low power mode) to an active mode, not while the computing device remains in the low power mode as recited in the ’330 Patent claims.

PO Resp. 13 (underlined text is in the PO Response)(citing Ex. 1003, 11:37–42; Ex. 2012 ¶¶ 64–68.

Thus, the key issue between the parties is *when* does Rosenberger exit its low power mode? Petitioner asserts Rosenberger exits its low power mode *after* the devices *exchange* weighted signal inputs and *after* the devices *determine* which device will respond to a user query, and thus anticipates claims 1, 9, and 17. Patent Owner asserts Rosenberger exits its low power mode *upon receiving a trigger word*, and thus does not anticipate these claims.

Because Rosenberger discloses several different embodiments, to some extent both parties are correct. *See* Reply 5 (asserting that Patent Owner’s arguments “inappropriately limit Rosenberger’s teachings as a whole to one specific example”); *id.* at 10–11. We agree with Petitioner, however, that at least one of the embodiments in Rosenberger discloses the claimed invention and thus anticipates claims 1, 9, and 17. Moreover, Patent Owner’s argument “relies on the erroneous assumption that the disclosure of multiple examples renders one example less anticipatory.” *Leggett & Platt, Inc. v. VUTEk, Inc.*, 537 F.3d 1349, 1356 (Fed. Cir. 2008).

Regarding the weighted signal, Rosenberger discloses

the control device of the invention uses one of several different coordination methods to achieve the design goal of avoiding duplicate responses to the same speech command while selecting a single control device best able to successfully interact with the system user. All of the methods involve the calculation and production by the control device of a numeric weighted signal (WS) that quantifies the “quality” of a just-recognized speech command.

Ex. 1003, 10:34–41. Rosenberger also discloses that “there are a number of methods for the devices to be coordinated so that only the single device with the highest weighted signal proceeds to interact with the user and/or process the speech command.” *Id.* at 11:27–31.

Rosenberger discloses that its device D starts out in a “low power ‘listening’ mode.” *Id.* at 8:17–20. Rosenberger then explains that “[u]pon receiving and recognizing a speech trigger phrase *and determining that it is in a better position to handle subsequent user interaction than any other device* that simultaneously recognized the same speech trigger phrase” (*id.* at 8:20–24), the winning device “beeps, changes its status light 34 or plays a prerecorded or synthesized audio message” to prompt the user to state its query. *Id.* at 8:30–32. This indicates that the determination based on the weighted signal is made while the system is still in low power mode and *before* the winning device “beeps, changes its status light 34 or plays a prerecorded or synthesized audio message” to prompt the user to state its query” (*id.*).

Petitioner’s Declarant, Dr. Johnson, testifies that “Rosenberger discloses that, *after* the ‘control devices’ exchange WS values as part of such ‘coordination,’ only the ‘control device’ with the highest WS value exits the claimed low power mode.” Ex. 1002 ¶ 131 (emphasis added)).

Dr. Johnson also explains the various alternative ways in which Rosenberger may function.

while Rosenberger discloses that each ‘control device’ is generally configured to exit *a low power mode* based on a user’s utterance that contains a ‘trigger phrase,’ Rosenberger further discloses instances in which ‘control devices’ receive a user’s utterance that contains a trigger phrase’ but do not immediately exit the *low power mode*. Rather, the ‘control devices’ engage in

a ‘coordination’ method that involves exchanging WS values, where the result of the ‘coordination’ method dictates what ‘control device is to exit the *low power mode*. Thus, in sum, Rosenberger discloses that (i) each ‘control device’ operates in the claimed *low power mode* while the applicable ‘coordination method’ is performed –which involves each ‘control device’ transmitting its own WS value to one or more other devices in the system– and (ii) only after the applicable ‘coordination method’ is performed, does a single ‘control device’ exit the claimed *low power mode* to interact with the user and/or process speech commands.

Ex. 1002 ¶ 114 (emphasis in original). Dr. Johnson concludes by testifying that “it is my opinion that Rosenberger discloses element 9.2.” *Id.* at 115.

Patent Owner focuses on the disclosure at column 11 in Rosenberger disclosing a “Broadcast and Monitor method” of determining which device has the highest weighted signal. PO Resp. 17 (citing Ex. 1003, 11:55–59); *see also* Ex. 1003, 11:37–40. Using this method, Rosenberger states “[f]irst, the microcontroller waits to be “triggered.” (i.e., switched from a ‘listening’ mode to an active mode)”. *Id.* at 11:40–42 (disclosing that “[f]irst, the microcontroller waits to be ‘triggered.’ (i.e., switched from a ‘listening’ mode to an active mode)”). Patent Owner also relies on the testimony of Mr. Lipoff. PO Resp. 22–27 (citing Ex. 2012 ¶¶ 71–74, 76–78).

Patent Owner appears to ignore, however, the disclosure of Rosenberger at column 8, which we repeat below. As disclosed in Rosenberger,

Preferably, device D [a control device] *normally operates in a low power ‘listening mode’ in which it listens for a hands-free trigger phrase* (e.g., “Hello Voice Control or “Home Automation”) to be spoken by the system user. *Upon receiving and recognizing a speech trigger phrase [Rosenberger’s label for a “Hotword”] and determining that it is in a better position to*

handle subsequent user interaction than any other device that simultaneously recognized the same speech trigger phrase (see device coordination discussion below), or when the user depresses the “Push to Talk” button 24 on the device, or in the event the device is *instructed to wake up and expect a subsequent speech command by a system controller*, the device beeps, *changes its status light 34* [see Figs. 1, 2B] or plays a prerecorded or synthesized audio message (e.g. “How may I help you?”) through speaker 19 and/or 20 to prompt the user to say one of the speech commands from a known vocabulary or grammar stored in the speech pattern memory 14. Upon recognizing a particular speech command (and optionally confirming with the user), the device produces and transmits either to a system controller or directly to the automated mechanisms (via its communications circuitry 18) a corresponding system control signal designed to instruct the appropriate automated mechanisms to carry out the subject of the speech command.

Ex. 1003, 8:17–39 (emphases added).

As disclosed, Rosenberger’s “low power listening mode” is for listening for a trigger phrase. *Id.* at 8:17–19. It is only *after* detecting the trigger phrase and “*upon . . . determining that it is in a better position to handle subsequent user interaction than any other device* that simultaneously recognized the same speech trigger phrase” does Rosenberger’s device D “*change[] its status light 34*” to indicate its new “wake up” status as being ready for a subsequent speech command. *Id.* at 8:20–27 (emphases added).

Accordingly, giving substantial weight to this disclosure in Rosenberger and the related testimony by Dr. Johnson, based on a preponderance of the evidence before us, we determine that Rosenberger discloses the elements and limitations of clause [9.2].

iv. Clause [9.3]

[9.3] while the computing device remains in the low power mode, receiving, by the computing device and from an additional computing device that is in a low power mode, an additional message;

Clause [9.2] focused on a single computing device *transmitting* a message while in a low power mode. Clause [9.3] is similar to clause [9.2] but recites that the first computing device also *receives* a message from “an additional computing device” that also is in a “low power mode.”

Petitioner asserts Rosenberger discloses the steps, elements, and limitations of clause [9.3]. Pet. 39–43 (citing Ex. 1002 ¶¶ 116–127).

Patent Owner asserts that “Petitioner’s analysis is flawed” because “as discussed with respect to [9.2], mere *capability* is insufficient to establish anticipation.” PO Resp. 27. We have addressed this argument in footnote 12 of this Decision.

Patent Owner also asserts that “Rosenberger’s device is not configured to receive an additional message while remaining in low power mode, and the hypothetical capabilities alleged by Petitioner are insufficient to establish anticipation.” *Id.* at 28 (citing Ex. 2012 ¶ 81). *Id.* at 29 (citing Ex. 2012 ¶¶ 81–83). We disagree.

As we have discussed in our discussion of clause [9.2] and of Rosenberger, Rosenberger discloses in Figures 7–10 and the related written descriptions of these figures “various system architectures in which a plurality of the control devices [four control devices are shown] of the invention either coordinate amongst themselves [Figs. 9, 10] or utilize a coordinating controller [Figs. 7, 8] to decide which control device will handle a given speech command when more than one control device has

been simultaneously triggered.” Ex. 1003, 5:64–6:3. As also explained above, triggering a device with a “trigger phrase” as disclosed in Rosenberg is not synonymous with exiting Rosenberg’s low power mode.

Accordingly, based on a preponderance of the evidence before us, we determine that Rosenberg discloses devices that both send and receive “messages” with each other while in a low power mode, which meets the requirements of clause [9.3] based on the same analysis as in clause [9.2] above.

v. Clause [9.4]

[9.4] based on the message and the additional message, determining, by the computing device, to exit the low power mode.

Essentially, clause [9.4] states that the final step of determining which of the computing devices will remain in the low power mode, and by implication, which one computing device, the one with the highest hotword confidence score, will exit the low power mode.

Petitioner asserts Rosenberg discloses the steps, elements, and limitations in clause [9.4]. Pet. 43–46 (citing Ex. 1002 ¶¶ 128–136). Petitioner cites, for example, Rosenberg’s disclosure that the control devices are “coordinated so that only the single device with the highest weighted signal proceeds to interact with the user and/or process the speech command.” Pet. 44 (quoting Ex. 1003, 11:27–31). All other control devices in Rosenberg remain in the low power mode. *Id.* at 45. Any other result would not eliminate the “duplicate response problem” that Rosenberg seeks to avoid, which occurs when several devices simultaneously recognize the same speech command or trigger phrase. *See* Ex. 1003, 3:30–35.

Patent Owner asserts that “Petitioner’s arguments for [9.4] suffer from the same deficiencies as [9.3].” PO Resp. 31. According to Patent Owner, Petitioner’s argument that “only the ‘control device’ with the highest WS [weighted signal] value exits the claimed *low power mode*” (Pet. 43–44 (italics in Petition)) is “incorrect.” *Id.* We have determined otherwise in our discussion of clause [9.3] and of Rosenberger above.

Accordingly, based on a preponderance of the evidence before us, we determine that Rosenberger discloses the steps, elements and limitations of clause [9.4].

b. Claims 1 and 17

The claim chart in Section II.D.2 of this Decision establishes the substantial similarity of claims 1, 9, and 17

Petitioner asserts elements 1.1, 1.2, 1.3, and 1.4 of claim 1 map to elements 9.1, 9.2, 9.3, and 9.4 of claim 9, respectively. Pet. 47. Petitioner concludes “Rosenberger discloses claim 1 for the same reasons it discloses claim 9. *Id.* (citing Ex. 1002 ¶¶ 137–139).

Petitioner also asserts elements 17.0, 17.1, 17.2, 17.3, and 17.4 of claim 17 map to elements 9.0, 9.1, 9.2, 9.3, and 9.4 of claim 9, respectively. *Id.* at 48–49. Petitioner concludes “Rosenberger discloses claim 17 for the same reasons it discloses claim 9.” *Id.* (citing Ex. 1002 ¶¶ 140–142).

Patent Owner also argues each clause in claims 1 and 17 with its corresponding clause in claim 9. *See, e.g.*, PO Resp. 10 (arguing collectively clauses [1.2], [9.2], [17.2]).

c. Conclusion for Claims 1, 9, 17

In *inter partes* reviews, a petitioner bears the burden of proving unpatentability of the challenged claims, and the burden of persuasion never

shifts to the patent owner. *Dynamic Drinkware, LLC v. Nat'l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015). To prevail in this proceeding, Petitioner must support its challenge by a preponderance of the evidence. 35 U.S.C. § 316(e); 37 C.F.R. § 42.1(d). Based on our analysis above of claims 1, 9, and 17, we determine that Petitioner has met its burden of proof for these claims. Accordingly, we determine claims 1, 9, and 17 are not patentable.

3. *Dependent claims*
2, 3, 5, 7, 10, 11, 13, 15, and 18

Claims 2, 10, and 18 are identical and depend on independent claims 1, 9, and 17, respectively.

Claims 3 and 11 are identical and depend on independent claims 1 and 9, respectively.

Claims 5 and 13 are identical and depend on independent claims 1 and 9, respectively.

Claims 7 and 15 are identical and depend on independent claims 1 and 9, respectively.

Petitioner provides a clause-by-clause analysis of where each element or limitation in dependent claims 2, 3, 5, 7, 10, 11, 13, 15, and 18 is disclosed in Rosenberger, and explains why and how Rosenberger anticipates these claims. Pet. 49–60. Throughout its analysis, Petitioner cites the declaration testimony of Dr. Johnson for evidentiary support.

Patent Owner does not address the patentability of dependent claims 2, 5, 7, 10, 13, 15, and 18. *See* PO Resp. 35 (addressing only dependent claims 3 and 11); Sur-Reply; *see also* Reply 17 (“Google does not dispute

that Rosenberger anticipates the additional limitations of dependent claims 2/10/18, 5/13, and 7/15.’”).

We adopt Petitioner’s undisputed findings and conclusions concerning dependent claims 2, 3, 5, 7, 10, 11, 13, 15, and 18. Based on these findings and conclusions, we determine that Petitioner has met its burden of proof for these claims. Accordingly, we determine claims 2, 3, 5, 7, 10, 11, 13, 15, and 18 are not patentable.

a. Dependent Claims 3 and 11

Claims 3 and 11 are identical and depend on independent claims 1 and 9, respectively. Claims 3 and 11 each state:

the computing device is configured to exit the low power mode in response to detecting an utterance of the particular, predefined hotword, and

the additional computing device is configured to exit the low power mode in response to detecting an utterance of the particular, predefined hotword.

E.g., Ex. 1001, 16:45–51 (for claim 3).

Petitioner states “Rosenberger discloses these additional elements.” Pet. 53 (citing Ex. 1002 ¶¶ 154–161).

According to Petitioner, a person of ordinary skill, which we have determined is a person with the equivalent of a four-year degree from an accredited institution (typically denoted as a B.S. degree) in electrical or computer engineering, or a related field, and approximately 2–4 years of experience in the field of voice and/or audio processing, “would have understood that Rosenberger’s ‘trigger phrase’ is synonymous with *the particular, predefined* hotword.” Pet. 53 (citing Ex. 1002 ¶ 156).

According to Petitioner,

Rosenberger discloses that, based on detecting a ‘trigger phrase’ within a voice utterance from a user, each ‘control device’ is configured to exit (e.g., wake up from) a *low power mode* in which the ‘control device’ declines to interact with the user and process voice commands after a ‘trigger phrase.’ *See, e.g.,* Ex.1003, 2:6-9 (‘When a **trigger phrase** is **detected**, the **system** ‘**wakes up**’ and prompts the user to say one or more commands from an expected command phrase vocabulary.’).

Pet. 54 (emphasis in Petition) (citing also Ex. 1003, 8:17–32).

Patent Owner asserts two reasons why claims 3 and 11 are patentable. First, Patent Owner asserts “Rosenberger cannot anticipate claims 3 and 11 because Rosenberger does not anticipate claims 1 and 9 from which they respectively depend.” PO Resp. 35 (citing Ex. 2012 ¶¶ 89–109).

We have determined that claims 1 and 9 are anticipated by Rosenberger and are not patentable. Thus, we disagree with Patent Owner’s first assertion.

Second, Patent Owner asserts “Rosenberger does not expressly disclose ‘detecting an utterance of the particular, predefined hotword.’” *Id.*

According to Patent Owner, “Petitioner’s reliance on the prior art in combination with Rosenberger’s disclosed invention cannot sustain its anticipation challenge.” *Id.* We disagree with Patent Owner’s premise and its conclusion. Moreover, Patent Owner does not direct us to any persuasive evidence to support this argument. We acknowledge that Petitioner refers to the prior art. For example, Petitioner states “[i]n line with Rosenberger’s discussion of prior art systems that utilize low-power modes and low-power communication technology (*id.*, 1:35-2:1), *Rosenberger explains* that an objective of the invention is to provide a ‘speech recognition device and system that operate using low-power....’ *Id.*, 3:45-49.” Pet. 15. Here,

Petitioner refers to the prior art and points out that Rosenberger's explanation of Rosenberger's disclosed invention is "in line with" the prior art. Petitioner is relying on Rosenberger's explanation, not the prior art, as the basis for the anticipation of the challenged claims.

Hotword Detection/Speech recognition

Patent Owner also provides an extensive argument regarding claims 3 and 11 asserting that "Hotword Detection Is Distinct from Speech Recognition." PO Resp. 36–47 (*see* Section heading V.D.2, PO Resp. 36).

Patent Owner argues the '330 Patent "distinguishes hotword detection from speech recognition and describes the use of a distinct *on-device*"¹³ hotword detector, *i.e.*, 'hotworder' to detect a hotword in incoming audio data." PO Resp. 36 (emphasis added) (citing Ex. 2012 ¶ 92); Sur-reply 14 ("Nothing in Rosenberger indicates that the control device 'detects' a trigger phrase (*i.e.*, via hotword detection as compared to speech recognition). We note, however, that contrary to Patent Owner's assertion, the '330 patent states it "performs *speech recognition techniques to determine whether the hotword was spoken* and, if so, awaits an ensuing command or query." Ex. 1001, 1:65–67 (emphasis added). As explained further in the '330 patent,

In an example environment, the hotword used to invoke the system's attention are the words "OK computer." Consequently,

¹³ The '330 patent does not claim an "on-device" hotword detector. Here, Patent Owner apparently confuses the claims of the '330 patent with the claims of the related '398 patent, challenged in IPR2023-00118, which do claim an "on-device" hotworder. The written descriptions in the '330 and '398 patents do not mention an "on-device" hotworder, although Figure 1 in each patent shows the hotworder as contained within computing devices 106, 108, and 110.

each time the words “OK computer” are spoken, it is picked up by a microphone, conveyed to *the system, which performs speech recognition techniques to determine whether the hotword was spoken* and, if so, awaits an ensuing command or query. Accordingly, utterances directed at the system take the general form [HOTWORD] [QUERY], where “HOTWORD” in this example is “OK computer” and “QUERY” can be any question, command, declaration, or other request that can be speech recognized, parsed and acted on by the system, either alone or in conjunction with the server via the network.

Id. at 1:61–2:6 (emphasis added). The ’330 patent does not define how “speech recognition techniques” differ from “speech recognition.” Here, the hotword is a two-word phrase. The system listens for this exact spoken phrase, and, if spoken, the system is then ready to listen to, apply speech recognition to, and act on a subsequent user query. Rosenberger does the same thing in the same way with its “trigger phrase.”

While we recognize the difference between listening for a spoken phrase and speech recognition, we provide substantial probative weight to the statement in the ’330 patent about the substantial similarity between “hotwords” and speech recognition. Both use “speech recognition techniques.” Ex. 1001, 1:63–2:6. Moreover, we have not been directed to any step, element, or limitation in the challenged claims requiring the argued distinctions between “hotwords” and speech recognition.

Based on the evidence and analysis above, we determine that Petitioner has met its burden of proof for claims 3 and 11. Accordingly, we determine claims 3 and 11 are not patentable.

*E. Ground 2 – Obviousness of Claims 4, 6, 12, and 14
in View of Rosenberger and Basye*

Petitioner asserts dependent claims 4, 6, 12, and 14 are unpatentable under § 103 asserting that these claims would have been obvious in view of Rosenberger and Basye. Pet. 61–70.

We first summarize the disclosure in Basye.

1. Basye (Ex. 1013)

We make the following findings concerning the disclosure in Basye. Basye is titled “Speech Recognition Power Management.” Ex. 1013, code (54). As summarized in its Abstract, Bayse discloses:

[p]ower consumption for a computing device may be managed by one or more keywords. For example, if an audio input obtained by the computing device includes a keyword, a network interface module and/or an application processing module of the computing device may be activated. The audio input may then be transmitted via the network interface module to a remote computing device. Such as a speech recognition server. Alternately, the computing device may be provided with a speech recognition engine configured to process the audio input for on-device speech recognition.

Id., code (57). Basye discloses using a speech processing module to “determine a score or confidence level whose value corresponds to a likelihood that a keyword is actually present in the speech” and “[i]f the score satisfies a threshold, the speech processing module” may determine that the keyword is present in the speech but “if the score does not satisfy the threshold, the speech processing module” may determine that there is no keyword in the speech. *Id.* ¶ 27. Bayse’s focus, however, is on power management for the disclosed devices. As disclosed in Bayse, “aspects of the present disclosure are directed to power management for speech

recognition.” *Id.* ¶ 10. A computing device may be provided with a “power management subsystem” that selectively activates or deactivates one or more modules of the computing device. Ex. 1013 ¶ 10. “This activation may be responsive to an audio input that includes one or more pre-designated spoken words,” which Bayse refers to as “keywords.” *Id.*

Bayse also discloses:

By selectively activating modules of the computing device, the power management subsystem may advantageously improve the energy efficiency of the computing device. The power management subsystem may further improve the energy efficiency of the computing device by selectively activating one or more of its own modules. While such implementations are particularly advantageous for computing devices that rely on battery power, all computing devices for which power management may be desirable can benefit from the principles of the present disclosure.

Id. ¶ 14.

2. *Claims 4, 12 and Claims 6, 14*

Claims 4 and 12 are identical and depend from claims 1 and 9, respectively. Claims 4 and 12 state:

determining, by the computing device, that the audio data includes an utterance of a particular, predefined hotword without performing automated speech recognition processing on the audio data.

E.g., Ex. 1001, 16:52–56 (for claim 4).

Claims 6 and 14 are identical and also depend from claims 1 and 9, respectively. Claims 6 and 14 state:

determining, by the computing device, a hotword score that reflects a likelihood that the audio data includes an utterance of the particular, predefined hotword; and

determining, by the computing device, that the hotword score satisfies a threshold,

wherein transmitting the message is based on determining that the hotword score satisfies the threshold.

Id. at 16:62–17:2 (for claim 6).

Petitioner provides a clause-by-clause analysis of where each element or limitation in dependent claims 4 and 12 is disclosed in Rosenberger or Basye. Pet. 62–66. Petitioner also explains the motivation and reasons why it would have been obvious to combine the disclosures from these two references with a reasonable expectation of success. *Id.* at 65–66.

Throughout its analysis, Petitioner cites the declaration testimony of Dr. Johnson for evidentiary support. Petitioner provides a similar analysis for claims 6 and 14. Pet. 67–70.

In its Response, Patent Owner asserts:

Arguments relating to dependent claims 4 and 6 are deficient for at least the same reasons discussed for independent claim 1 from which they depend. Arguments relating to dependent claims 12 and 14 are also deficient for at least the same reasons discussed for independent claim 9 from which they depend.

PO Resp. 48.

We did not agree with Patent Owner’s arguments regarding the asserted “deficiencies” of Rosenberger against the independent claims, and determine these same arguments are equally unpersuasive against dependent claims 4, 6, 12, and 14. Patent Owner does not otherwise substantively address claims 6 and 14 in its Response or its Sur-reply.

Regarding claims 4 and 12, Patent Owner asserts:

It would not have been obvious to modify Rosenberger in accordance with the teachings of Basye to not perform speech recognition on the audio data that corresponds to an utterance of

the particular, predefined hotword, (Pet. 61–66), as recited, because the entire premise of Rosenberger’s invention is to perform speech recognition.

PO Resp. 48 (citing Ex. 2012 ¶ 113). Patent Owner also states “Rosenberger’s disclosure is replete with passages evidencing an express need and preference for speech recognition.” *Id.* at 49. Petitioner then lists ten “bullet points” where Rosenberger discusses “speech recognition.” *Id.* at 49–51. Patent Owner concludes that “Petitioner’s proposed modification, *which removes speech recognition processing of the trigger phrase*, upends . . . the entire premise of Rosenberger.” *Id.* at 52 (citing Ex. 2012 ¶¶ 111–116). We disagree.

Rosenberger’s extensive reference to “speech recognition” is consistent with the technology disclosed and claimed in Rosenberger, and in the ’330 patent. We note that the ’330 patent uses the term “speech recognition” 81 times. In describing its “Technical Field” the ’330 patent states “[t]his specification generally relates to systems and techniques for recognizing the words that a person is speaking, otherwise referred to as *speech recognition*.” Ex. 1001, 1:17–19 (emphasis added). Thus, a reference that anticipates the invention claimed in the ’330 patent likely also would emphasize speech recognition.

Petitioner provides three arguments as to why a person of ordinary skill in the art would have been motivated to combine Rosenberger with Basye. Pet. 61–62.

Petitioner first argues that “Rosenberger and Basye are in the same field of endeavor and relate to voice-controllable computing devices.” *Id.* at 61 (citing Ex. 1003, 1:8–13 (“The present invention relates . . . to improvements in speech-recognizing control devices that respond to oral

commands from a system user to produce a desired effect or result.”); Ex. 1013 ¶ 1 (“Computing devices may include speech recognition capabilities.... A computing device may also be capable of processing the recognized speech for specific speech recognition applications.”)); *see* Ex. 1002 ¶ 182).

Second, Petitioner argues that “Rosenberger and Basye both disclose techniques for intelligently ‘waking’ a voice-controllable computing device using a ‘trigger phase’ (or ‘keyword’ or ‘wakeword’ in Basye’s terminology).” Pet. 61–62; *see* Ex. 1002 ¶ 183.

Third, Petitioner argues that “Rosenberger discloses that any ‘control device’ could be a ‘tabletop device’ that ideally is powered by one or more batteries.” Pet. 62 (citing Ex. 1013, 8:55–67, 9:37–40); *see* Ex. 1002 ¶ 194. In support of these arguments, Dr. Johnson testifies that “a POSITA seeking to implement one of Rosenberger’s battery-powered ‘control devices’ would have been motivated ... to look to teachings regarding intelligent power management to improve energy efficiency, such as the teachings of Basye.” Ex. 1002 ¶ 184.

Patent Owner asserts:

it would not have been obvious to modify Rosenberger in accordance with the teachings of Basye to not perform speech recognition on the audio data that corresponds to an utterance of the particular, predefined hotword (Pet. 69-73), as recited, because the entire premise of Rosenberger’s invention is to perform speech recognition.

PO Resp. 47 (citing Ex. 2012 ¶ 111). According to Patent Owner, a person of ordinary skill “would not have been motivated to modify Rosenberger such that its computing device no longer performs speech recognition on the audio data corresponding to the trigger phrase (alleged ‘hotword’) because

the proposed modification alters the speech recognition principles forming the entire premise of Rosenberger.” PO Resp. 51 (citing 2012 ¶¶ 109–114).

As we have noted above, the summary of the invention disclosed in the ’330 patent states “each time the words ‘OK computer’ [the chosen “hotword”] are spoken, it is picked up by a microphone, conveyed to the system, *which performs speech recognition techniques to determine whether the hotword was spoken and, if so, awaits an ensuing command or query.* Ex. 1001, 1:62–66. Thus, before the “the output of processing the audio data” is transmitted, the system first uses “speech recognition techniques” to determine whether the words spoken are, in fact, the adopted “hotword.” The ’330 patent does not define how “speech recognition techniques” differ from “speech recognition.” To the extent there is a substantive difference, Petitioner relies on Basye.

As Petitioner explains, using Basye instead of Rosenberger’s speech recognition, would “help reduce ‘false positive’ detections of a ‘trigger phrase’ (or ‘wakeword’).” Pet. 70 (citing Ex. 1002 ¶ 214). Petitioner also asserts that “Basye explains that the use of ‘thresholding’ can help reduce false positives, which is a situation where a computing device incorrectly determines that speech is present or that speech includes a wakeword.” *Id.* (citing Ex. 1013 ¶¶ 51, 74–75). This provides a good reason why a person of ordinary skill would want to modify, and improve, Rosenberger’s performance.

Based on our analysis above of claims 4, 6, 12, and 14, we determine that Petitioner has met its burden of proof for these claims. Accordingly, we determine claims 4, 6, 12, and 14 are not patentable.

III. CONCLUSION¹⁴

Based on the evidence and analysis above, we determine that Petitioner has established that all challenged claims, claims 1–7, 9–15, 17, 18 are not patentable based the asserted grounds.

IV. ORDER

In consideration of the foregoing, it is hereby:

ORDERED that, that Petitioner has shown by a preponderance of the evidence that claims 1–7, 9–15, 17, 18 are unpatentable;

FURTHER ORDERED that, because this is a Final Written Decision, the parties to the proceeding seeking judicial review of the decision must comply with the notice and service requirements of 37 C.F.R. § 90.2.

¹⁴ Should Patent Owner wish to pursue amendment of the challenged claims in a reissue or reexamination proceeding subsequent to the issuance of this decision, we draw Patent Owner's attention to the April 2019 *Notice Regarding Options for Amendments by Patent Owner Through Reissue or Reexamination During a Pending AIA Trial Proceeding*. See 84 Fed. Reg. 16,654 (Apr. 22, 2019). If Patent Owner chooses to file a reissue application or a request for reexamination of the challenged patent, we remind Patent Owner of its continuing obligation to notify the Board of any such related matters in updated mandatory notices. See 37 C.F.R. § 42.8(a)(3), (b)(2).

V. SUMMARY TABLE

Claims	35 U.S.C. §	Reference(s)/Basis	Claims Shown Unpatentable	Claims Not shown Unpatentable
1–3, 5, 7, 9– 11, 13, 15, 17, 18	102	Rosenberger	1–3, 5, 7, 9– 11, 13, 15, 17, 18	
4, 6, 12, 14	103	Rosenberger and Basye	4, 6, 12, 14	
Overall Outcome			1–7, 9–15, 17, 18	

For PETITIONER:

Michael P. Boyea
Jae Y. Pak
Cole B. Richter
LEE SULLIVAN SHEA & SMITH LLP
boyea@ls3ip.com
pak@ls3ip.com
richter@ls3ip.com

For PATENT OWNER:

Erika Harmon Arner
Cory C. Bell
Alissa E. Green
Misook Kim
Safiya Anguilar
FINNEGAN, HENDERSON, FARABOW,
GARRETT, & DUNNER LLP
erika.arner@finnegan.com
cory.bell@finnegan.com
alissa.green@finnegan.com
misook.kim@finnegan.com
safiya.anguilar@finnegan.com