

UNITED STATES INTERNATIONAL TRADE COMMISSION

Washington, D.C.

In the Matter of

**CERTAIN AUDIO PLAYERS AND
COMPONENTS THEREOF (II)**

Inv. No. 337-TA-1330

ORDER NO. 14: CONSTRUING TERMS OF THE ASSERTED CLAIMS

(February 15, 2023)

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	RELEVANT LAW	1
III.	TECHNICAL BACKGROUND AND LEVEL OF ORDINARY SKILL.....	7
IV.	THE ASSERTED PATENTS	9
A.	748 Patent.....	9
B.	128 Patent.....	13
C.	311 Patent.....	17
D.	615 Patent.....	19
V.	CLAIM CONSTRUCTION.....	22
A.	748 Patent - “predetermined spoken audio command”.....	24
B.	748 patent - “[perform/performing] speech recognition”	26
C.	748 Patent - “[determine/determining] . . . based at least in part on information received from the one or more additional computing devices, whether the first computing device should perform speech recognition on the spoken audio input”	27
D.	748 Patent - “a command to temporarily refrain from performing speech recognition”	29
E.	128 Patent - “other devices of the plurality of electronic devices” / “quality scores generated by the other devices of the plurality of electronic devices”	31
	311 Patent - “second values . . . from other electronic devices” / “second values from other electronic devices of the plurality of electronic devices”	31
F.	128 Patent - “forgo outputting an audible response”	37
	128 Patent - “forgoing outputting a response”	37
	311 Patent - “[forgoing / forgo] responding”	37
G.	128 Patent - “in accordance with a determination that the quality score generated by the first electronic device is the highest . . . in accordance with a determination that the quality score generated by the first electronic device is not the highest”	42
	311 Patent - “in accordance with a determination that the first value is higher than the second values . . . in accordance with a determination that the first value is not higher than the second scores”	42

H.	128 Patent - “cause the electronic device”	46
I.	311 Patent - “the second scores”	48
J.	311 Patent - “the detected output”	49
K.	615 Patent - “short-range signal transmission path”	50
L.	615 Patent - “in response to receiving the request, receiving one or more network configuration packets”	54

I. INTRODUCTION

This investigation was instituted by the Commission on September 15, 2022 to determine whether audio player devices—specifically, voice controllable audio player devices and networked audio player devices—infringe U.S. Patent Nos. 9,632,748 (“the 748 patent”), 9,812,128 (“the 128 patent”), 11,024,311 (“the 311 patent”), and 11,050,615 (“the 615 patent”) (altogether, “the Asserted Patents”). *See* 87 Fed. Reg. 56701 (Sep. 15, 2022). The complainant is Google LLC (“Google”) and the respondent is Sonos, Inc. (“Sonos”).

Pursuant to the procedural schedule, the parties filed a joint claim construction chart setting forth certain terms to be construed, and initial and rebuttal claim construction briefs. On December 21, 2022, a previously scheduled Markman hearing was cancelled. Order No. 8. On January 10, 2023, the parties filed a final updated joint claim chart.¹ EDIS Doc. ID 787658.

II. RELEVANT LAW

“An infringement analysis entails two steps. The first step is determining the meaning and scope of the patent claims asserted to be infringed. The second step is comparing the properly construed claims to the device accused of infringing.” *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 976 (Fed. Cir. 1995) (en banc) (internal citations omitted), *aff’d*, 517 U.S. 370 (1996). Claim construction is a “matter of law exclusively for the court.” *Id.* at 970-71. “The construction of claims is simply a way of elaborating the normally terse claim language in order to understand

¹ The briefs and amended chart submitted by the parties are hereafter referred to as:

CIMB	Google’s Initial <i>Markman</i> Brief
CRMB	Google’s Rebuttal <i>Markman</i> Brief
RIMB	Sonos’ Initial <i>Markman</i> Brief
RRMB	Sonos’ Rebuttal <i>Markman</i> Brief
JC	Updated Joint Claim Construction Chart

and explain, but not to change, the scope of the claims.” *Embrex, Inc. v. Serv. Eng’g Corp.*, 216 F.3d 1343, 1347 (Fed. Cir. 2000).

Claim construction focuses on the intrinsic evidence, which consists of the claims themselves, the specification, and the prosecution history. *See Phillips v. AWH Corp.*, 415 F.3d 1303, 1314 (Fed. Cir. 2005) (en banc); *see also Markman*, 52 F.3d at 979. As the Federal Circuit in *Phillips* explained, courts must analyze each of these components to determine the “ordinary and customary meaning of a claim term” as understood by a person of ordinary skill in the art at the time of the invention. 415 F.3d at 1313. “Such intrinsic evidence is the most significant source of the legally operative meaning of disputed claim language.” *Bell Atl. Network Servs., Inc. v. Covad Commc’ns Grp., Inc.*, 262 F.3d 1258, 1267 (Fed. Cir. 2001).

“It is a ‘bedrock principle’ of patent law that ‘the claims of a patent define the invention to which the patentee is entitled the right to exclude.’” *Phillips*, 415 F.3d at 1312 (quoting *Innova/Pure Water, Inc. v. Safari Water Filtration Sys., Inc.*, 381 F.3d 1111, 1115 (Fed. Cir. 2004)). “Quite apart from the written description and the prosecution history, the claims themselves provide substantial guidance as to the meaning of particular claims terms.” *Id.* at 1314; *see also Interactive Gift Express, Inc. v. Compuserve Inc.*, 256 F.3d 1323, 1331 (Fed. Cir. 2001) (“In construing claims, the analytical focus must begin and remain centered on the language of the claims themselves, for it is that language that the patentee chose to use to ‘particularly point [] out and distinctly claim [] the subject matter which the patentee regards as his invention.’”). The context in which a term is used in an asserted claim can be “highly instructive.” *Phillips*, 415 F.3d at 1314. Additionally, other claims in the same patent, asserted or unasserted, may also provide guidance as to the meaning of a claim term. *Id.* “Courts do not

rewrite claims; instead, we give effect to the terms chosen by the patentee.” *K-2 Corp. v. Salomon S.A.*, 191 F.3d 1356, 1364 (Fed. Cir. 1999).

The specification “is always highly relevant to the claim construction analysis. Usually it is dispositive; it is the single best guide to the meaning of a disputed term.” *Phillips*, 415 F.3d at 1315 (quoting *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996)). “[T]he specification may reveal a special definition given to a claim term by the patentee that differs from the meaning it would otherwise possess. In such cases, the inventor’s lexicography governs.” *Id.* at 1316. “In other cases, the specification may reveal an intentional disclaimer, or disavowal, of claim scope by the inventor.” *Id.* As a general rule, however, the particular examples or embodiments discussed in the specification are not to be read into the claims as limitations. *Id.* at 1323; *Kamstrup A/S v. Axioma Metering UAB*, 43 F.4th 1374, 1383-4 (Fed. Cir. 2022) (refusing to read several limitations from embodiments into claims); *Sunoco Partners Marketing & Terminals L.P. v. U.S. Venture, Inc.*, 32 F.4th 1161, 1176 (Fed. Cir. 2022) (holding plain language of claim was nonetheless broader than specification’s “preferably” statement). In the end, “[t]he construction that stays true to the claim language and most naturally aligns with the patent’s description of the invention will be . . . the correct construction.” *Id.* at 1316 (quoting *Renishaw PLC v. Marposs Societa' per Azioni*, 158 F.3d 1243, 1250 (Fed. Cir. 1998)).

In addition to the claims and the specification, the prosecution history should be examined, if in evidence. *Phillips* at 1317; see *Liebel-Flarsheim Co. v. Medrad, Inc.*, 358 F.3d 898, 913 (Fed. Cir. 2004). The prosecution history can “often inform the meaning of the claim language by demonstrating how the inventor understood the invention and whether the inventor limited the invention in the course of prosecution, making the claim scope narrower than it would otherwise be.” *Phillips*, 415 F.3d at 1317; see *Chimie v. PPG Indus. Inc.*, 402 F.3d 1371, 1384 (Fed. Cir.

2005) (“The purpose of consulting the prosecution history in construing a claim is to exclude any interpretation that was disclaimed during prosecution.”).

When the intrinsic evidence does not establish the meaning of a claim, then extrinsic evidence (*i.e.*, all evidence external to the patent and the prosecution history, including dictionaries, inventor testimony, expert testimony, and learned treatises) may be considered. *Phillips*, 415 F.3d at 1317. Extrinsic evidence is generally viewed as less reliable than the patent itself and its prosecution history in determining how to define claim terms. *Id.* “The court may receive extrinsic evidence to educate itself about the invention and the relevant technology, but the court may not use extrinsic evidence to arrive at a claim construction that is clearly at odds with the construction mandated by the intrinsic evidence.” *Elkay Mfg. Co. v. Ebco Mfg. Co.*, 192 F.3d 973, 977 (Fed. Cir. 1999).

If, after a review of the intrinsic and extrinsic evidence, a claim term remains ambiguous, the claim should be construed so as to maintain its validity. *Phillips*, 415 F.3d at 1327. Claims, however, cannot be judicially rewritten in order to fulfill the axiom of preserving their validity. *See Rhine v. Casio, Inc.*, 183 F.3d 1342, 1345 (Fed. Cir. 1999). Thus, “if the only claim construction that is consistent with the claim’s language and the written description renders the claim invalid, then the axiom does not apply and the claim is simply invalid.” *Id.*

The construction of a claim term is generally guided by its ordinary meaning. However, courts may deviate from the ordinary meaning when: (1) “the intrinsic evidence shows that the patentee distinguished that term from prior art on the basis of a particular embodiment, expressly disclaimed subject matter, or described a particular embodiment as important to the invention”; or (2) “the patentee acted as his own lexicographer and clearly set forth a definition of the disputed claim term in either the specification or prosecution history.” *Edwards Lifesciences LLC v. Cook*

Inc., 582 F.3d 1322, 1329 (Fed. Cir. 2009); *see also GE Lighting Sols., LLC v. AgiLight, Inc.*, 750 F.3d 1304, 1309 (Fed. Cir. 2014) (“the specification and prosecution history only compel departure from the plain meaning in two instances: lexicography and disavowal.”); *Omega Eng’g, Inc. v. Raytek Corp.*, 334 F.3d 1314, 1324 (Fed. Cir. 2003) (“[W]here the patentee has unequivocally disavowed a certain meaning to obtain his patent, the doctrine of prosecution disclaimer attaches and narrows the ordinary meaning of the claim congruent with the scope of the surrender.”); *Rheox, Inc. v. Entact, Inc.*, 276 F.3d 1319, 1325 (Fed. Cir. 2002) (“The prosecution history limits the interpretation of claim terms so as to exclude any interpretation that was disclaimed during prosecution.”). Nevertheless, there is a “heavy presumption that a claim term carries its ordinary and customary meaning.” *CCS Fitness, Inc. v. Brunswick Corp.*, 288 F.3d 1359, 1366 (Fed. Cir. 2002) (citations omitted). The standard for deviating from the plain and ordinary meaning is “exacting” and requires “a clear and unmistakable disclaimer.” *Thorner v. Sony Computer Entm’t Am. LLC*, 669 F.3d 1362, 1366-67 (Fed. Cir. 2012); *see Epistar Corp. v. Int’l Trade Comm’n*, 566 F.3d 1321, 1334 (Fed. Cir. 2009) (requiring “expressions of manifest exclusion or restriction, representing a clear disavowal of claim scope” to deviate from the ordinary meaning) (citation omitted). As the Federal Circuit has explained, “[w]e do not read limitations from the specification into claims; we do not redefine words. Only the patentee can do that.” *Thorner*, 669 F.3d at 1366; *see Continental Circuits LLC v. Intel Corp.*, 915 F.3d 788, 797 (Fed. Cir. 2019); *Unwired Planet, LLC v. Apple Inc.*, 829 F.3d 1353, 1359 (Fed. Cir. 2016) (“[W]e have repeatedly held that it is ‘not enough that the only embodiments, or all of the embodiments, contain a particular limitation’ to limit claims beyond their plain meaning.”).

Courts are not required to construe every claim limitation of an asserted patent. *See O2 Micro Intern. Ltd. v. Beyond Innovation Technology Co.*, 521 F.3d 1351, 1362 (Fed. Cir. 2008)

(citations omitted); *Vanderlande Indus. Nederland BV v. Int’l Trade Comm’n*, 366 F.3d 1311, 1323 (Fed. Cir. 2004) (noting that the administrative law judge need only construe disputed claim terms). Rather, “claim construction is a matter of resolution of disputed meanings and technical scope, to clarify and when necessary to explain what the patentee covered by the claims.” *Id.* at 1362 (quoting *U.S. Surgical Corp. v. Ethicon, Inc.*, 103 F.3d 1554, 1568 (Fed. Cir. 1997)); *see also Embrex*, 216 F.3d at 1347 (“The construction of claims is simply a way of elaborating the normally terse claim language in order to understand and explain, but not to change, the scope of the claims.”) (citation omitted). In addition, “[a] determination that a claim term ‘needs no construction’ or has the ‘plain and ordinary meaning’ may be inadequate when a term has more than one ‘ordinary’ meaning or when reliance on a term’s ‘ordinary’ meaning does not resolve the parties’ dispute.” *O2 Micro*, 521 F.3d at 1361. Claim construction, however, is not an “obligatory exercise in redundancy.” *U.S. Surgical Corp.*, 103 F.3d at 1568. “[M]erely rephrasing or paraphrasing the plain language of a claim by substituting synonyms does not represent genuine claim construction.” *C.R. Bard, Inc. v. U.S. Surgical Corp.*, 388 F.3d 858, 863 (Fed. Cir. 2004). However, “there are limits to the court’s duties at the claim construction stage. For example, courts should not resolve questions that do not go to claim scope, but instead go to infringement . . . or improper attorney argument.” *Eon Corp. IP Holdings v. Silver Spring Networks*, 815 F.3d 1314, 1319 (Fed. Cir. 2016).

A claim must also be definite. Pursuant to pre-AIA 35 U.S.C. § 112(2): “The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.” 35 U.S.C. § 112(2). In *Nautilus, Inc. v. Biosig Instruments, Inc.*, 134 S. Ct. 2120, 2129 (2014), the Supreme Court held that § 112(2) requires “that a patent’s claims, viewed in light of the specification and prosecution

history inform those skilled in the art about the scope of the invention with reasonable certainty.” A claim is required to “provide objective boundaries for those of skill in the art,” and a claim term is indefinite if it “might mean several different things and no informed and confident choice is among the contending definitions.” *Interval Licensing LLC v. AOL, Inc.*, 766 F.3d 1364, 1371 (Fed. Cir. 2014). A patent claim that is indefinite is invalid. 35 U.S.C. § 282(b)(3)(A). “But a claim is not indefinite just because it is broad.” *Niazi Licensing Corp. v. St. Jude Medical SC*, 30 F.4th 1339, 1347 (Fed. Cir. 2022) (citing *BASF Corp. v. Johnson Matthey Inc.*, 875 F.3d 1360, 1367 (Fed. Cir. 2017)).

III. TECHNICAL BACKGROUND AND LEVEL OF ORDINARY SKILL

The Asserted Patents cover computerized devices linked together in a network. The claims of the 748 patent encompass methods and apparatuses for determining which devices among a group should perform speech recognition on a given spoken audio command. *See* 748 patent at cls. 1, 7. The claims of the 128 patent also encompass methods and apparatuses for the same determination, this time based on a “quality score.” *See* 128 patent at cls. 1, 6, 11. The 311 patent functions similarly, claiming methods and apparatuses for the same determination, but based on “first values” and “second values.” *See* 311 patent at cls. 1, 10, 16. The 615 patent claims are distinct, leaving aside the concept of speech recognition to encompass methods and apparatuses for commissioning electronic devices into a network. *See* 615 patent at cls. 1, 11, 16.

The parties are in near agreement as to the level of ordinary skill in the art for the Asserted Patents. CIMB at 8 (citing CIMB, Ex. 10 at 2, 4). Google proposes two definitions: one for the 748, 128, and 311 patents, and one for the 615 patent. *Id.* at 8-9. As to the former grouping, Google writes, “[t]he 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.” *Id.* As to the 615 patent, Google states, “at least (a) a Bachelor’s degree in computer science, computer engineering, electrical engineering, or an equivalent thereof, and (b) 2-4 years of professional experience in the fields of networking and network-based systems or applications or an equivalent level of skill, knowledge, and experience.”

Id. at 9. Sonos, on the other hand, proposes an amalgamative definition for all patents:

[The] equivalent of a four-year degree from an accredited institution (typically denoted as a B.S. degree) in computer science, computer engineering, electrical engineering, or an equivalent thereof, and approximately 2-4 years of experience in the field of voice and/or audio processing, networking and network based systems or applications, or an equivalent level of skill, knowledge, and experience.

RIMB at 4 n.4.

Google’s approach is generally more persuasive. Unlike the others, the 615 patent has little to do with audio capture, recognition, or playback. Its focus is computerized networking procedures. *See, e.g.*, 615 patent at Abstract, cls. 1-10. Thus, it makes little sense to define a person of ordinary skill in its field as having education and experience in “voice and/or audio processing.” The generic nature of all the patents’ specifications otherwise supports the lower number of years of education and/or experience present in all proposals.

Accordingly, a POSITA for the 748, 128, and 311 patents would have 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. A POSITA for the 615 patent would have a B.S. degree in electrical or computer engineering, computer science, or a related field, and approximately 2-4 years of experience in networking and network-based systems or applications or an equivalent level of skill, knowledge, and experience. In both cases experience may substitute for education and vice versa.

IV. THE ASSERTED PATENTS

The 748 patent, entitled “Device Designation for Audio Input Monitoring,” issued on April 25, 2017 to Alexander Faaborg. The 748 patent reports an assignment on its face to Google Inc. and claims priority to a provisional application filed June 24, 2014.

The 128 patent, entitled “Device Leadership Negotiation Among Voice Interface Devices,” issued on November 7, 2017 to Kenneth Mixter, Diego Melendo Casado, Alexander Houston Gruenstein, Terry Tai, Christopher Thaddeus Hughes, and Matthew Nirvan Sharifi. The 128 patent reports an assignment on its face to Google Inc. and claims priority to a provisional application filed October 9, 2014.

The 311 patent, entitled “Device Leadership Negotiation Among Voice Interface Devices,” issued on June 1, 2021 to Kenneth Mixter, Diego Melendo Casado, Alexander Houston Gruenstein, Terry Tai, Christopher Thaddeus Hughes, and Matthew Nirvan Sharifi. The 311 patent reports an assignment on its face to Google LLC and claims priority to a provisional application filed October 9, 2014.

The 615 patent, entitled “Apparatus and Method for Seamless Commissioning of Wireless Devices,” issued on June 29, 2021 to Jeffrey P. Matthews and Lee R. Taylor. The 615 patent reports an assignment on its face to Google LLC and claims priority to a provisional application filed April 11, 2012.

A. 748 Patent

The 748 patent has 20 claims. As of the date of this order, claims 1-4, 7, 9-12, 14, and 15 are asserted in this investigation. The asserted claims read as follows (with the first instance of the disputed terms highlighted in bold and agreed terms in italics):

1. A method comprising:

receiving, by a first computing device of a plurality of computing devices in physical proximity to one another, a spoken audio input that is also received by one or more additional computing devices of the plurality of computing devices;

determining, by the first computing device and **based at least in part on information received from the one or more additional computing devices, whether the first computing device should perform speech recognition on the spoken audio input**; and

responsive to determining that the first computing device should perform speech recognition on the spoken audio input:

sending, by the first computing device and to at least a second computing device from the plurality of computing devices, **a command to temporarily refrain from performing speech recognition**;

performing, by the first computing device, speech recognition on the spoken audio input;

determining, by the first computing device, based on performing speech recognition on the spoken audio input, whether the spoken audio input includes a **predetermined spoken audio command**; and

responsive to determining that the spoken audio input includes the predetermined spoken audio command, executing, by the first computing device, an action associated with the predetermined spoken audio command.

2. The method of claim 1, further comprising:

determining, by the first computing device, a current context of the first computing device,

wherein determining whether the first computing device should **perform speech recognition** on the spoken audio input is further based on the current context of the first computing device.

3. The method of claim 2, wherein determining the current context of the first computing device comprises determining one or more of: a location of the first computing device, a current time as defined by the first computing device, one or more applications installed at the *sec-end first computing device*, one or more applications currently executing at the first computing device, one or more networks available to the first computing device, one or more other computing devices in proximity to the first computing device, an operating mode of the first computing device, an ambient temperature of the location of the first computing device, an ambient noise level of the location of the first computing device, an ambient light level of the location of the first computing device, a movement of the first computing device, a name of a user of the first computing device, a user identification (UID) of the user of the first computing device, a social media

network service account associated with the user of the first computing device, one or more calendars associated with the user, or one or more social relationships of the user of the first computing device.

4. The method of claim 1,

wherein the information received from the one or more additional computing devices includes an indication of an audio input quality associated with a second computing device from one or more additional computing devices, and

wherein determining whether the first computing device should **perform speech recognition** on the spoken audio input is further based on the audio input quality associated with the second computing device and an audio input quality associated with the first computing device.

....

7. A first computing device comprising:

at least one processor; and

at least one module operable by the at least one processor to:

receive a spoken audio input that is also received by one or more additional computing devices in physical proximity to the first computing device;

determine, based at least in part on information received from the one or more additional computing devices, whether the first computing device should perform speech recognition on the spoken audio input; and

responsive to determining that the first computing device should perform speech recognition on the spoken audio input:

send, to at least a second computing device from a plurality of computing devices, **a command to temporarily refrain from performing speech recognition;**

perform speech recognition on the spoken audio input;

determine, based on performing speech recognition on the spoken audio input, whether the spoken audio input includes a **predetermined spoken audio command;** and

responsive to determining that the spoken audio input includes the predetermined spoken audio command, execute an action associated with the predetermined spoken audio command.

....

9. The computing device of claim 7,

wherein the at least one module is further operable by the at least one processor to determine a current context of the first computing device, and

wherein the at least one module is operable by the at least one processor to determine whether the first computing device should **perform speech recognition** on the spoken audio input further based on the current context of the first computing device.

10. The computing device of claim 9, wherein the at least one module operable to determine the current context of the first computing device is operable by the at least one processor to determine one or more of: a location of the first computing device, a current time as defined by the first computing device, one or more applications installed at the first computing device, one or more applications currently executing at the first computing device, one or more networks available to the first computing device, one or more other computing devices in proximity to the first computing device, an operating mode of the first computing device, an ambient temperature of the location of the first computing device, an ambient noise level of the location of the first computing device, an ambient light level of the location of the first computing device, a movement of the first computing device, a name of a user of the first computing device, a user identification (UID) of the user of the first computing device, a social media network service account associated with the user of the first computing device, one or more calendars associated with the user of the first computing device, or one or more social relationships of the user of the first computing device.

11. A non-transitory computer-readable storage medium encoded with instructions that, when executed, cause at least one processor of a first computing device to:

receive a spoken audio input that is also received by one or more additional computing devices in physical proximity to the first computing device;

determine, based at least in part on information received from the one or more additional computing devices, whether the first computing device should perform speech recognition on the spoken audio input; and

responsive to determining that the first computing device should perform speech recognition on the spoken audio input:

send, to at least a second computing device from a plurality of computing devices, **a command to temporarily refrain from performing speech recognition;**

perform speech recognition on the spoken audio input;

determine, based on performing speech recognition on the spoken audio input, whether the first audio input includes a **predetermined spoken audio command;** and

responsive to determining that the spoken audio input includes the predetermined spoken audio command, execute an action associated with the predetermined spoken audio command.

12. The non-transitory computer-readable storage medium of claim 11, further encoded with instructions that, when executed, cause the at least one processor to:

determine a current context of the first computing device,

wherein the instructions cause the at least one processor to determine whether the first computing device should **perform speech recognition** on the spoken audio input further based on the current context of the first computing device.

....

14. The non-transitory computer-readable storage medium of claim 11,

wherein the information received from the one or more additional computing devices includes an indication of audio input quality associated with a second computing device from the one or more additional computing devices, and

wherein the instructions cause the at least one processor to determine whether the first computing device should **perform speech recognition** on the spoken audio input further based on the audio input quality associated with the second computing device and an audio input quality associated with the first computing device.

15. The first computing device of claim 7,

wherein the information received from the one or more additional computing devices includes an indication of an audio input quality associated with a second computing device from the one or more additional computing devices, and

wherein the at least one module is operable by the at least one processor to determine whether the first computing device should **perform speech recognition** on the spoken audio input further based at least in part on the audio input quality associated with the second computing device and an audio input quality associated with the first computing device.

B. 128 Patent

The 128 patent has 15 claims. As of the date of this order, claims 1-15 are asserted in this investigation. The asserted claims read as follows (with the first instance of the disputed terms highlighted in bold and agreed terms in italics):

1. A method, comprising:

at a first electronic device of a plurality of electronic devices, each electronic device of the plurality of electronic devices comprising one or more microphones, a speaker, one or more processors, and memory storing one or more programs for execution by the one or more processors:

detecting a voice input;

determining a quality score for the detected voice input;

communicating the quality score to the **other devices of the plurality of electronic devices**;

receiving **quality scores generated by the other devices of the plurality of electronic devices** for detection of the voice input by the other devices;

in accordance with a determination that the quality score generated by the first electronic device is the highest amongst the generated quality score and received quality scores for the voice input, outputting an audible and/or a visual response to the detected voice input, wherein the other devices of the plurality of electronic devices **forgo outputting an audible response** to the detected voice input; and

in accordance with a determination that the quality score generated by the first electronic device is not the highest amongst the quality scores for the voice input generated by the plurality of electronic devices, **forgoing outputting a response** to the detected voice input.

2. The method of claim 1, wherein the plurality of electronic devices is communicatively coupled through a local network; and

the communicating and receiving are performed through the local network.

3. The method of claim 1, wherein the quality score comprises a confidence level of detection of the voice input.

4. The method of claim 1, wherein the quality score comprises a signal-to-noise rating of detection of the voice input.

5. The method of claim 1, further comprising:

recognizing a command in the voice input; and

in accordance with a determination that a type of the command is related to the first electronic device, outputting an audible and/or a visual response to the detected voice input.

6. A first electronic device of a plurality of electronic devices, *each of the plurality of electronic devices comprising*:

one or more microphones;

a speaker;

one or more processors; and

memory storing *one or more programs to be executed by the one or more processors, the one or more programs comprising instructions for:*

detecting a voice input;

determining a quality score for the detected voice input;

communicating the quality score to the **other devices of the plurality of electronic devices**;

receiving **quality scores generated by the other devices of the plurality of electronic devices** for detection of the voice input by the other devices;

in accordance with a determination that the quality score generated by the first electronic device is the highest amongst the generated quality score and received quality scores for the voice input, outputting an audible and/or a visual response to the detected voice input, wherein the other devices of the plurality of electronic devices **forgo outputting an audible response** to the detected voice input; and

in accordance with a determination that the quality score generated by the first electronic device is not the highest amongst the quality scores for the voice input generated by the plurality of electronic devices, **forgoing outputting a response** to the detected voice input.

7. The device of claim 6, wherein the plurality of electronic devices is communicatively coupled through a local network; and

the communicating and receiving are performed through the local network.

8. The device of claim 6, wherein the quality score comprises a confidence level of detection of the voice input.

9. The device of claim 6, wherein the quality score comprises a signal-to-noise rating of detection of the voice input.

10. The device of claim 6, further comprising instructions for:

recognizing a command in the voice input; and

in accordance with a determination that a type of the command is related to the first electronic device, outputting an audible and/or a visual response to the detected voice input.

11. A non-transitory computer readable storage medium storing one or more programs, the one or more programs comprising instructions, which, when executed by a first electronic device of a plurality of electronic devices, each of the plurality of electronic devices comprising one or more microphones, a speaker, and one or more processors, **cause the electronic device** to perform operations comprising:

detecting a voice input;

determining a quality score for the detected voice input;

communicating the quality score to the **other devices of the plurality of electronic devices**;

receiving **quality scores generated by the other devices of the plurality of electronic devices** for detection of the voice input by the other devices;

in accordance with a determination that the quality score generated by the first electronic device is the highest amongst the generated quality score and received quality scores for the voice input, outputting an audible and/or a visual response to the detected voice input, wherein the other devices of the plurality of electronic devices **forgo outputting an audible response** to the detected voice input; and

in accordance with a determination that the quality score generated by the first electronic device is not the highest amongst the quality scores for the voice input generated by the plurality of electronic devices, **forgoing outputting a response** to the detected voice input.

12. The computer readable storage medium of claim 11, wherein the plurality of electronic devices is communicatively coupled through a local network; and

the communicating and receiving are performed through the local network.

13. The computer readable storage medium of claim 11, wherein the quality score comprises a confidence level of detection of the voice input.

14. The computer readable storage medium of claim 11, wherein the quality score comprises a signal-to-noise rating of detection of the voice input.

15. The computer readable storage medium of claim 11, further comprising instructions which, when executed by the electronic device, **cause the electronic device** to perform operations comprising:

recognizing a command in the voice input; and

in accordance with a determination that a type of the command is related to the first electronic device, outputting an audible and/or a visual response to the detected voice input.

C. 311 Patent

The 311 patent has 20 claims. As of the date of this order, claims 1-3, 8-12, and 14-20 are asserted in this investigation. The asserted claims read as follows (with the first instance of the disputed terms highlighted in bold):

1. A method performed at a first electronic device of a plurality of electronic devices, each of the plurality of electronic devices comprising one or more microphones, one or more processors, and memory storing one or more programs for execution by the one or more processors, the method comprising:

detecting a voice input from a user;

determining a first value associated with the voice input;

receiving **second values** associated with the voice input **from other electronic devices** of the plurality of electronic devices; and

in accordance with a determination that the first value is higher than the second values, responding to the detected input.

2. The method of claim 1, further comprising:

in accordance with a determination that the first value is not higher than the second scores, forgoing responding to the detected input.

3. The method of claim 2, wherein **forgoing responding** to the detected input further comprises changing a state of operation of the first electronic device.

....

8. The method of claim 1, wherein responding to the detected input comprises outputting an audible and/or a visual response to the detected voice input.

9. The method of claim 1, wherein the first electronic device further comprises a speaker, and responding to the detected input further comprises outputting an audible response.

10. A first electronic device of a plurality of electronic devices, the first electronic device comprising:

one or more microphones;

one or more processors; and

memory storing one or more programs to be executed by the one or more processors, the one or more programs comprising instructions for:

detecting a voice input from a user;

determining a first value associated with the voice input;

receiving **second values** associated with the voice input **from other electronic devices** of the plurality of electronic devices; and

in accordance with a determination that the first value is higher than the second values, responding to the detected input.

11. The first electronic device of claim 10, the one or more programs further comprising instructions for:

recognizing a command in the voice input; and

in accordance with a determination that a type of the command is related to the first electronic device, responding to the detected voice input.

12. The first electronic device of claim 10, the one or more programs further comprising instructions for:

communicating to the other devices of the plurality of electronic devices to **forgo responding to the detected output**.

....

14. The first electronic device of claim 10, wherein:

the plurality of electronic devices is communicatively coupled through a local network; and

the receiving is performed through the local network.

15. The first electronic device of claim 10, wherein the first electronic device is configured to be awakened by a plurality of affordances including a voice-based affordance; and

wherein detecting the voice input is in response to the awakening of the first electronic device.

16. A system comprising a plurality of electronic devices, each of the electronic devices including one or more microphones, one or more processors, and memory storing one or more programs for execution by the one or more processors, each of the electronic devices programmed to:

detect a voice input;

determine a first value associated with the detected voice input;

receive respective **second values from other electronic devices of the plurality of electronic devices**;

compare the first value and the respective second values; and

in accordance with a determination that the first value is higher than the respective second values, respond to the voice input.

17. The system of claim 16, wherein the first value comprises a first quality score of the voice input.

18. The system of claim 17, wherein the first quality score comprises a confidence level of detection of the voice input.

19. The system of claim 17, wherein the first quality score comprises a signal-to-noise rating of detection of the voice input.

20. The system of claim 16, wherein each of the respective second values is indicative of a quality of the voice input detected by a respective one of the other electronic devices.

D. 615 Patent

The 615 patent has 20 claims. As of the date of this order, claims 1-3, 5-12, and 15-19 are asserted in this investigation. The asserted claims read as follows (with the first instance of the disputed terms highlighted in bold):

1. A method for commissioning a target device onto a communication network, comprising:

at the target device including a microprocessor, a wireless radio, and commissioning logic:

receiving a request to initiate a commissioning process with the target device;

establishing a **short-range signal transmission path** with a wireless device independently of the communication network;

in response to receiving the request, receiving one or more network configuration packets from the wireless device via the short-range signal transmission path, wherein the one or more network configuration packets convey network configuration data associated with the communication network according

to a predetermined data format, the network configuration data including at least a security passphrase for accessing the communication network;

executing an application on the target device to decode and process the one or more network configuration packets and recover the network configuration data; and

after receiving the network configuration packets, employing via the wireless radio the network configuration data conveyed in the one or more network configuration packets to enable the target device to join the communication network.

2. The method of claim 1, wherein the wireless device is selected from a group consisting of a smartphone and a tablet computer.

3. The method of claim 1, wherein information received via the **short-range signal transmission path** is received at a first wireless radio of the target device, and the first wireless radio employs at least one of a plurality of radio modulation techniques including frequency shift keying (FSK), phase shift keying (PSK), on-off keying (OOK), and quadrature amplitude modulation (QAM).

....

5. The method of claim 1, wherein the **short-range signal transmission path** is associated with a communication channel based on one of Bluetooth communication, radio frequency identification (RFID), Zigbee, radio frequency for consumer electronics (RF4CE), Z-Wave, and near field communication.

6. The method of claim 1, further comprising:

contacting the wireless device via the communication network to indicate that the target device has successfully entered the communication network.

7. The method of claim 6, further comprising:

after contacting the wireless device, continuing the commission process, including receiving, via the communication network, device configuration data from the wireless device.

8. The method of claim 7, wherein the device configuration data including at least one of an Internet Protocol address, proxy, and gateway configuration.

9. The method of claim 1, further comprising:

generating one of a visual indication and an audio indication to a user about a status of the commissioning process.

10. The method of claim 1, further comprising an application that includes a sequence of program instructions directed towards commissioning the target device onto the communication network.

11. A target device, comprising:

a microprocessor;

a wireless radio;

a commissioning logic; and

a memory storing programs for execution by the target device, the programs further including instructions for:

receiving a request to initiate a commissioning process with the target device;

establishing a **short-range signal transmission path** with a wireless device independently of a communication network;

in response to receiving the request, receiving one or more network configuration packets from the wireless device via the short-range signal transmission path, wherein the one or more network configuration packets convey network configuration data associated with the communication network according to a predetermined data format, the network configuration data including at least a security passphrase for accessing the communication network;

executing an application on the target device to decode and process the one or more network configuration packets and recover the network configuration data; and

after receiving the network configuration packets, employing via the wireless radio the network configuration data conveyed in the one or more network configuration packets to enable the target device to join the communication network.

12. The target device of claim 11, further comprising an application that includes a sequence of program instructions directed towards commissioning the target device onto the communication network.

....

15. The target device of claim 11, wherein the communication network includes a wireless local area network (WLAN).

16. A non-transitory computer readable storage medium storing one or more programs configured for execution by a target device for commissioning the target device onto a communication network, the target device including a microprocessor, a wireless radio, and commissioning logic, the one or more programs comprising instructions for:

receiving a request to initiate a commissioning process with the target device;

establishing a **short-range signal transmission path** with a wireless device independently of the communication network;

in response to receiving the request, receiving one or more network configuration packets from the wireless device via the short-range signal transmission path, wherein the one or more network configuration packets convey network configuration data associated with the communication network according to a predetermined data format, the network configuration data including at least a security passphrase for accessing the communication network;

executing an application on the target device to decode and process the one or more network configuration packets and recover the network configuration data; and

after receiving the network configuration packets, employing via the wireless radio the network configuration data conveyed in the one or more network configuration packets to enable the target device to join the communication network.

17. The non-transitory computer readable storage medium of claim 16, wherein the one or more programs further comprise instructions for:

generating one of a visual indication and an audio indication to a user about a status of the commissioning process.

18. The non-transitory computer readable storage medium of claim 16, further comprising an application that includes a sequence of program instructions directed towards commissioning the target device onto the communication network.

19. The non-transitory computer readable storage medium of claim 16, wherein the one or more programs further comprise instructions for:

contacting the wireless device via the communication network to indicate that the target device has successfully entered the communication network.

V. CLAIM CONSTRUCTION

The disputed claim terms are summarized in the parties' Joint Claim Construction Chart.

See JC. They are discussed individually below.

The parties also provide two agreed constructions, one for the 748 patent and one for the 128 patent:

Term	Agreed Construction
"the sec-end first computing device" (748 patent at cl. 3)	"the first computing device"

“each of the plurality of electronic devices comprising . . . one or more programs to be executed by the one or more processors, the one or more programs comprising instructions for” (128 patent at cl. 6)	Each device of the “plurality of electronic devices” includes the recited “memory storing one or more programs to be executed by the one or more processors,” and the “one or more programs” comprises instructions for performing the operations recited in the remainder of the claim
---	--

JC at 1.

The construction for the term of the 128 patent is accepted. It is consistent with the plain and ordinary meaning and intrinsic evidence, even if it would appear to be largely redundant to the actual claim language.

The construction for the term of the 748 patent is not accepted. It appears to be aimed at correcting the patentee’s errant inclusion of the non-word “sec-end.” There has not been sufficient briefing as to whether this is proper under the law, however; either as a correction of an error or otherwise understandable, and thus forgivable, lack of antecedent basis. *See, e.g., Novo Indus., L.P. v. Micro Molds Corp.*, 350 F.3d 1348, 1357 (Fed. Cir. 2003) (district courts may “correct obvious minor typographical and clerical errors” in patents, but not “major errors”); *see, e.g., Certain Variable Speed Wind Turbines and Components Thereof*, Inv. No. 337-TA-641, Comm’n Op. at 36 (Mar. 2, 2010) (“*Wind Turbines*”) (“The Commission lacks the authority to correct inventorship under Section 256 or any other statutory provision, and the Commission’s authority in this regard must be conferred by statute.”); *In re Downing*, 754 F.App’x 988, 990, 996 (Fed. Cir. 2018) (holding that a lack of antecedent basis need not result in indefiniteness when the claim appraises a skilled artisan of its scope, and observing that “claim 1’s recitation of one end user could only refer to the end user using the product”).

A. 748 Patent - “predetermined spoken audio command”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“any audio input usable to initiate interaction with a computing device, such as a spoken hotword”

JC at 1.

For “predetermined spoken audio command,” the parties dispute what is meant by “command.” Google takes the position that no construction is necessary and that “initiate interaction” as in Sonos’ construction vitiates the ordinary meaning, which is “a directive that can be executed by the device” (CIMB at 17 (citing CIMB, Ex. 6 at ¶¶ 67-68)), and in addition, improperly replaces “command” with “spoken hotword,” which is too narrow (*id.* at 19). Google argues, “[t]he specification consistently and repeatedly associates predetermined audio commands with corresponding actions” (*id.* at 18 (citing 748 patent at 1:15-18, 4:8-10, 4:14-21, 9:61-62, 20:15-18, 24:36-39, 24:41-46); CRMB at 2); and when the specification uses “spoken hotword” it prefaces that characterization with “generally, may be” (CRMB a 1 (citing 748 patent at 3:67)). Google contends this permissive language is not “exacting” so as to trigger inventor lexicography. *Id.* at 1-2 (citing *Thorner*, 669 F.3d at 1365).

Sonos argues that inventor lexicography has occurred and points to the following excerpt: “[a] predetermined audio command, generally, may be any audio input (*e.g.*, a word, a phrase, etc.) [that] is usable to initiate interaction with a computing device.” RIMB at 9-10 (citing 748 patent at 3:67-4:2); RRMB at 1. Sonos also claims “predetermined spoken audio command” “is [] a coined phrase that has no plain and ordinary meaning” such that “the ALJ’s construction for this term *must* be limited to the disclosure in the specification, which supports Sonos’s proposed construction.” *Id.* at 10 (emphasis added) (citing *Iridescent Networks, Inc. v. AT&T Mobility, LLC*, 933 F.3d 1345, 1351-53 (Fed. Cir. 2019); *Goldenberg v. Cytogen, Inc.*, 373 F.3d 1158, 1164

(Fed. Cir. 2004)). Sonos rejects Google’s criticism and urges that its construction “broadly covers ‘any audio input usable to initial interaction with a computing device.’” RRMB at 2 (emphasis in original). Sonos views Google as attempting to exclude “hotwords” from the scope of the term. *See id.* at 2-3.

Sonos is less persuasive than Google and so its construction is not accepted. The specification does not support the lexicography which would be required for Sonos’ “initiate interaction” definition. As Google observes, the “generally, may be” phrase communicates that an express definition is not being given (748 patent at 3:67-4:2) and the specification uses a simpler term—“predetermined audio command”—in basic ways that connote a plain and ordinary meaning:

One or more of audio input modules 6 may associate a predetermined audio command with an action to be performed by the respective computing device. For instance, one or more of audio input modules 6 may associate a word or phrase (e.g., “Hello phone,” “Okay car,” etc.) with entering an audio input mode, thereby allowing the user to provide further audio input. Other examples of predetermined audio inputs and associated actions may include a word or phrase (e.g., “answer”) that causes a computing device to answer a telephone call, a word or phrase (e.g., “navigation,” “directions,” etc.) that causes a computing device to provide directions, a word or phrase (e.g., “send email,” “send text message,” etc.) that causes a computing device to send a message, and others. In some examples, predetermined audio commands may be user-configurable. That is, a user may create, modify, enable, disable, or discard various predetermined audio commands. In some examples, which predetermined audio commands a computing device is configured to act on may be determined (e.g., by the computing device) based on a number of other criteria, such as the applications installed at the computing device, the applications executing at the computing device, the time of day, or other criteria.

Id. at 4:8-30. This conclusion is further supported by Google’s un rebutted expert declaration. CIMB, Ex. 6 at ¶¶ 67-68. And Sonos concedes (reasonably) that there is no pertinent difference between “predetermined audio command” and “predetermined spoken audio command.” RIMB at 10, n.5. Moreover, Sonos’ proposed “initiate interaction” construction cannot be correct, because

the claims require that interaction already be initiated by “performing . . . speech recognition on the spoken audio input” before determining whether that input includes the “predetermined spoken audio command.” 748 patent at cl. 1; *see also id.* at cls. 7, 11.

Accordingly, Sonos’ proposed construction of “predetermined spoken audio command” is rejected, which resolves the parties’ dispute, so the term is not otherwise construed.

B. 748 patent - “[perform/performing] speech recognition”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning), i.e., “transforming speech to corresponding text”	“processing spoken audio to determine whether it includes a predetermined spoken audio command”

JC at 1.

For “[perform/performing] speech recognition,” Sonos’ construction is not persuasive. The plain and ordinary meaning of the phrase is as the 748 patent’s specification describes it—the process of transforming spoken audio into text (*i.e.*, to transcribe) or some other form readable by the claimed computing device:

For instance, audio input modules may be operable to utilize speech recognition techniques to determine text from obtained audio input, and compare the text to a set of predetermined audio commands (e.g., stored at the respective computing device or other location) to determine whether the obtained audio input includes a predetermined audio command.

748 patent at 4:57-59. Google presents un rebutted expert opinion to the same effect (CIMB, Ex. 6 at ¶¶ 52-58) in addition to other supporting extrinsic evidence (*see, e.g.*, CIMB, Ex. 11 at 3).

Sonos’ construction, on the other hand, sweeps in the next claim element, which concerns what happens subsequent to the “recognition”—the analysis of the produced text—and so it is incorrect. Sonos refers to excerpts from various “other” Google patents yet makes no assertion that these are related to the 748 patent, and they are not. *See* RIMB at 11-12 (citing RIMB, Exs. A, B, C, D). They therefore do not constitute intrinsic evidence. Moreover, the excerpts do not

define “speech recognition” for what it is but instead describe the ends it might be used for. *See* RIMB, Ex. A at 7:12-17 (“keyword detector module 124 may perform speech recognition on the audio signal 122 to identify a portion of the signal corresponding to the keyword”); RIMB, Ex. B at 1:14-20; RIMB, Ex. C at 1:62-66; RIMB, Ex. D at 1:63-67. Sonos then cites a district court case, involving unrelated parties and unrelated patents (from a much earlier timeframe), where “speech recognition” was construed. RIMB at 12 (citing *Phoenix Sols., Inc. v. DirecTV Grp., Inc.*, No. CV08-0984-MRPSSX, 2008 WL 5727419, at *15 (C.D. Cal. Nov. 4, 2008)). For these reasons the case is of dubious value, but even then it settled on a construction that favors Google, not Sonos: “[t]he Court therefore construes ‘speech recognition operations’ to be ‘processing tasks performed on a digitized speech utterance to recognize symbols contained in the speech utterance.’” *Id.* The aforementioned “symbols” would be characters, as in text, and not necessarily “spoken commands.” 2008 WL 5727419, at *4 n.3. Sonos’ remaining point, that Google’s construction, which stops at turning speech into text, somehow “exclude[s] ‘hotword detection’ from ‘speech recognition,’” (*see* RIMB at 4-6) is irrelevant. As stated, the analysis of the text to determine if hotword, keyword, or predetermined commands are present occurs subsequent to the speech recognition.

Accordingly, Sonos’ proposed construction of “[perform/performing] speech recognition” is rejected, which resolves the parties’ dispute, so the term is not otherwise construed.

C. 748 Patent - “[determine/determining] . . . based at least in part on information received from the one or more additional computing devices, whether the first computing device should perform speech recognition on the spoken audio input”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“determining whether the first computing device should perform speech recognition on the spoken audio input without any of the plurality of computing devices processing the

	spoken audio input to determine whether it includes a predetermined audio command”
--	--

JC at 1.

For “[determine/determining] . . . based at least in part on information received from the one or more additional computing devices, whether the first computing device should perform speech recognition on the spoken audio input,” Sonos again presents an unpersuasive construction. To the basic concept of determining whether a computing device should perform speech recognition, Sonos adds “without any of the plurality of computing devices processing the spoken audio input to determine whether it includes a predetermined audio command.” There is nothing in the meaning of the words of the claim limitation that could be interpreted in this way, so Sonos reaches into subsequent claim language to read in this condition:

Specifically, the language in all the claims of the ’748 Patent recites that a determination is made as to “whether the first computing device should perform speech recognition on the spoken audio input,” and if so, the other computing devices should “refrain from performing speech recognition.” As such, the determination as to “whether the first computing device should perform speech recognition on the spoken audio input” must be done without any of the plurality of computing devices performing speech recognition, *i.e.*, “without any of the plurality of computing devices processing the spoken audio input to determine whether it includes a predetermined audio command.”

RIMB at 12-13. Even then, Sonos misrepresents the language of the claim to make its point. The claim does not recite that the “other computing devices should ‘refrain from performing speech recognition.’” It only states that the first computing device, after having determined that it should be the one to perform speech recognition, “send[s] . . . to at least a second computing device from the plurality of computing devices, a command to temporarily refrain from performing speech recognition.” 748 patent at cl. 1. Whether or not those other computing devices had been performing speech recognition (and thus need to stop) is outside the scope of this language.

Sonos’ discussion of the specification is equally unavailing. Sonos seems to reason that the specification fails to disclose determining if the first device should perform speech recognition based on the spoken audio input—and instead discloses a variety of other criteria—and therefore that a negative limitation should be read into the claim term. *See* RIMB at 14. But if every concept that a specification did not disclose was then read into the claims as a negative limitation, there would be no end to the exercise. And the passages Sonos highlights in no way rise to the level of exacting disclaimer required under *Thorner*. *See id.* at 14-15 (citing 748 patent at 7:37-8:13, 8:26-52, 25:22-67, Fig. 5); RRMB at 7-8 (citing same); 669 F.3d at 1366-67.

Accordingly, no intrinsic evidence supports adding the negative limitation advanced by Sonos. So it is not accepted and the parties’ dispute is resolved.

D. 748 Patent - “a command to temporarily refrain from performing speech recognition”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“a message instructing the second computing device to temporarily cease performing any and all speech recognition”

JC at 1.

The parties’ dispute for “a command to temporarily refrain from performing speech recognition” is primarily over the “any and all” phrase in Sonos’ construction. Sonos argues “restraining the performance of some, but not all, speech recognition would not be ‘refrain[ing] from performing speech recognition,’ and would fail to meet the desired goals expressed in the ‘748 Patent.” RIMB at 15. Sonos reasons, “[a]ny other interpretation for the language used in the is claim term would be illogical, and thus, inappropriate.” *Id.* (citing *KEYnetik, Inc. v. Samsung Elecs. Co.*, 837 F. App’x 786, 793 (Fed. Cir. 2020)). For support, Sonos points to a variety of

specification passages which discuss the benefits from ceasing unnecessary speech recognition activities. *Id.* at 16-17 (citing 748 patent at 2:52-63, 3:7-13, 5:31-59, 10:7-12, Figs. 4, 5).

Sonos is persuasive here. The plain and ordinary meaning of a command to refrain from an activity means to cease that activity, for any purpose or reason. Doing so would accomplish the 748 patent's stated goals of avoiding wasted processing time and energy, as Sonos notes. RIMB at 16 (citing 748 patent at 3:7-13, 5:31-39, 10:7-12). And Google does not so much dispute the point as express confusion as to the meaning of "any and all." CIMB at 16 ("The plain meaning of this claim language . . . the second computing device should not perform speech recognition for a limited amount of time."), 17 ("It is not clear how, if at all, Sonos intends ['any and all'] to affect the claim scope"); CRMB at 8 ("It is not clear how Sonos' construction, adding 'any and all' to the claim, differs in any meaningful way from the plain language or would clarify the claims' scope for a POSITA."), 9 (stating the invention's benefits are realized "when only one device monitors for and/or processes audio input"). As noted, "any and all" clarifies that the command must not simply direct the second computing device from performing *some amount* of speech recognition, but all speech recognition.

This interpretation is supported by the language of claim 1 itself. The preceding limitation explicitly introduces method steps which must be done in response to "spoken audio input." 748 patent at cl. 1. The three subsequent limitations are also explicit that their recited actions ("performing," "determining," or "executing") are done on or to that same "spoken audio input." *Id.* ("performing . . . speech recognition on the spoken audio input," "determining . . . whether the spoken audio input includes . . .," "executing . . . an action associated with the predetermined spoken audio command [as included within the spoken audio input]."). The limitation at issue, however, differs in that its action of "refrain[ing]" is not linked to the "spoken audio input"; *i.e.*,

the claim does not recite “a command to temporarily refrain from performing speech recognition on the spoken audio input.” See *id.* (emphasized language added). Thus, a POSITA would understand “refrain from performing speech recognition” to mean any speech recognition and not just recognition of the “spoken audio input.”

The remainder of Sonos’ construction, however, addressing the target of the command and replacing “refrain” to “cease,” without explanation, is not appropriate and not adopted. *Hoganas AB v. Dresser Indus., Inc.*, 9 F.3d 948, 950 (Fed. Cir. 1994) (“It is improper for a court to add ‘extraneous’ limitations to a claim, that is, limitations added ‘wholly apart from any need to interpret what the patentee meant by particular words or phrases in the claim.’”); *C.R. Bard*, 388 F.3d at 863 (“[M]erely rephrasing or paraphrasing the plain language of a claim by substituting synonyms does not represent genuine claim construction.”).

Accordingly, “a command to temporarily refrain from performing speech recognition” is construed as “a command to temporarily refrain from performing any and all speech recognition.”

E. 128 Patent - “other devices of the plurality of electronic devices” / “quality scores generated by the other devices of the plurality of electronic devices”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“other devices” means “two or more other devices”; and “quality scores” means “two or more quality scores”

JC at 2.

311 Patent - “second values . . . from other electronic devices” / “second values from other electronic devices of the plurality of electronic devices”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“second values” means “two or more second values”; and “electronic devices” means “two or more electronic devices”

JC at 2.

Within the 128 and 311 patents are claim terms that involve a “plurality of electronic devices” and, within that group, a “first electronic device” and then “other electronic devices of the plurality of electronic devices.” 128 patent at cl. 1; 311 patent at cl. 1. The parties dispute how many devices “other electronic devices” requires—whether it can be satisfied by a single “other” device, or must be two or more. The dispute is the same for “quality scores” (128 patent at cl. 1), “second values” (311 patent at cl. 1), and “other electronic devices” (311 patent at cl. 1).

Google argues each of these terms can be satisfied by a single device, score, or value. CIMB at 26, 29. For support, Google looks to preceding claim language introducing “a plurality of electronic devices” which ordinarily means two or more. *Id.* at 26. Google then reasons that two total electronic devices must be sufficient and so “[a] POSITA would understand that, with the claimed plurality of electronic devices including the ‘first electronic device,’ there need be only one ‘other device’ and one ‘quality score’ received from that ‘other device’ to constitute a ‘plurality.’” *Id.* (citing CIMB, Ex. 6 at ¶ 79). Google asserts this comes from a need for “grammatical consistency”:

In the ’128 patent, the claims use the plural form of “other devices” and “quality scores” to maintain grammatical consistency with the “plurality of other devices.” Given that context, and the patent’s disclosure of a “plurality of electronic devices” and a “first electronic device,” the recited “other devices” and “quality scores” should be construed to encompass a configuration with one other device and its associated quality score (e.g., a two-device “plurality of electronic devices”).

Id. (citations omitted); *see id.* at 27-28 (citing *Dayco Prods., Inc. v. Total Containment, Inc.*, 258 F.3d 1317, 1328 (Fed. Cir. 2001)), 28 (citing 128 patent at 28:63-29:5). Lastly, Google views the case cited by Sonos, *Apple Inc. v. MPH Techs. Oy*, 28 F.4th 254 (Fed. Cir. 2022) (“*MPH Techs.*”), as inapposite because that case lacked specification language which is present in the 128 patent—

language “mak[ing] clear that the claims encompass embodiments of a ‘plurality of electronic devices’ in which there is only a single ‘other device,’ ‘second value,’ or ‘quality score.’” CRMB at 10.

Sonos contends “[a]ssigning a meaning of ‘two or more’ to plural nouns is ‘simply an application of the general rule that claim terms are usually given their plain and ordinary meaning.’” RIMB at 18 (citing *MPH Techs.*, 28 F.4th at 261-62; *Leggett & Platt, Inc. v. Hickory Springs Mfg. Co.*, 285 F.3d 1353, 1357 (Fed. Cir. 2002)). Sonos suggests the 128 and 311 patents’ “decision to recite plural nouns as opposed to singular nouns was deliberate” in light of the reappearance of these pluralized words within the claims, and the distinct use of “one or more” for other claim elements such as processors and programs. *Id.* at 19 (citing 128 patent at cls. 1, 6, 11; 311 patent at cl. 16), 20 (citing 128 patent at cl. 1; 311 patent at cl. 1). Sonos reasons, “if the Applicant truly intended for ‘quality scores’ and ‘other devices’ to be singular, it would have done the same for these terms and placed the “one or more” in front of those terms itself.” *Id.* at 20 (citing *Helmsderfer v. Bobrick Washroom Equip., Inc.*, 527 F.3d 1379, 1381-82 (Fed. Cir. 2008); *Medtronic, Inc. v. Guidant Corp.*, 465 F.3d 1360, 1377 (Fed. Cir. 2006)); RRMB at 9 (“Its choice not to recite ‘one or more’ here is controlling.”). For further support, Sonos points to an embodiment in the 311 patent involving three devices (311 patent at Fig. 4A, 20:55-21:35) and reasons:

Given this disclosure, the Applicant was thus within its rights to expressly claim this embodiment. Indeed, this is the exact configuration called for by the claims when they recite “a first electronic device,” which receives “quality scores generated by the other devices of the plurality of electronic devices.” Thus, the minimum number of devices called for by the claims is three devices.

Id. at 21. And in its reply brief, Sonos views Google as improperly “reading in the language ‘one or more’ as a prefix to every plural noun (*i.e.*, ‘[one or more] quality scores generated by the [one or more] other devices’).” RRMB at 9.

The issue is a close one but Sonos is not persuasive. While the pluralized version of a noun normally indicates more than one, in the context of the inventions of the 128 and 311 patents, the claimed “other devices” or “other electronic devices” may be met by a single additional device apart from the “first electronic device.” Essentially, there must be a first device and then any number of additional “other” devices. This ordinary grammatical technique is evident elsewhere in patents, including dependent claim 7 of the 311 patent reciting “voice inputs”:

The method of claim 1, further comprising: at the first electronic device: detecting subsequent voice inputs from the user; pre-processing the subsequent voice inputs; and transmitting the pre-processed voice inputs to a server system wherein the other electronic devices forgo the detecting, pre-processing, and transmitting.

311 patent at cl. 7. While “voice inputs” is plural, normally indicating at least two inputs, in context it is clear this claim would cover detecting, pre-processing, and transmitting just one “subsequent” input.

And contrary to Sonos’ assertion (RIMB at 21 (“Nowhere does the specification use strictly plural nouns to mean both plural and nonplural”)), the specification does use plurals in places where it is understood that only a single item would suffice:

A location (e.g., a room or space within a home) may include multiple devices that include voice assistant systems and one or more users of such devices. As the devices are more than ever configured to be capable of accessing and presenting user-specific information, it is desirable for a voice assistant device to be able to handle multiple users gracefully and be able to provide user-specific information to the appropriate user. It is also desirable for there to be a leader amongst the voice assistant devices that is responsible for responding to the *user’s voice inputs*, in order to reduce user confusion. Further, it is desirable for a voice

assistant device to include *noise mitigation measures* in case the *voice assistant devices* are located at a noisy location.

311 patent at 1:60-2:5 (emphasis added). As for Sonos' highlighted three-device embodiment (RIMB at 21), there is also an explicit teaching of just two:

FIG. 6 is a flow diagram illustrating a method 1200 of negotiating device leadership among multiple voice interface devices, in accordance with some implementations. The method 1200 is implemented at an electronic device (e.g., device 1002-1, 1002-2, or 1002-3) of *two or more* electronic devices 190 (electronic devices 1002), where each of the two or more electronic devices includes one or more microphones, a speaker, one or more processors and memory storing one or more programs for execution by the one or more processors.

128 patent at 28:63-29:5; 311 patent at 29:8-17. There is no more reason to think that claim 1 of each patent is describing the three-device embodiment than this two-device embodiment. It should also be noted that the patents fail to disclose any differences in operation or importance to three devices in the plurality as opposed to two, and Sonos certainly does not allude to any. *See generally* RIMB at 17-22; RRMB at 9-12. This further diminishes the likelihood a POSITA would read the claim to be so limited.

With that said, the patents' frequent use of "one or more" throughout the specification and within the claims for other elements does support Sonos' position. *See, e.g.*, RRMB at 9 (citing *Allen Eng'g Corp. v. Bartell Indus., Inc.*, 299 F.3d 1336, 1349 (Fed. Cir. 2022)). Indeed, the specification also uses the parenthetical "(s)" to cover situations where one or more than one item would be acceptable. *See, e.g.*, 128 patent at 11:35-37 ("It will be appreciated that in some implementations the electronic device(s) 190 also function in an environment without smart home devices."). These techniques could also have been used for "communicating the quality score to the other devices of the plurality of electronic devices" or "receiving second values associated with the voice input from other electronic devices of the plurality of electronic devices." RIMB at

19-20 (“if the Applicant truly intended for ‘quality scores’ and ‘other devices’ to be singular, it would have done the same for these terms and placed the ‘one or more’ in front of those terms itself.”). However, this would have resulted in the loss of an important technical objective which the patent discusses at length; namely, that only one device out of the group act as the leader and process voice inputs. *See, e.g.*, 128 patent at 22:66-23:45 (describing “leadership determination process” and “just the leader device 1002-1 outputs the response per the leadership determination”); 311 patent at 23:9-55 (same). With “one or more” inserted, the claim could be met after receiving only one “value” or “score” even though others exist. This would allow for two leaders, even though Sonos acknowledges there should only be one, stating “[i]ndeed, the point of the invention is to have *multiple* devices carry out the leader election process in order to select one of them as the leader.” RRMB at 19 (emphasis in original). This is not a preferred outcome in claim construction:

Based on the full teaching of the specification, we conclude that Wi-Fi's proposed construction of claim 15 is unreasonable. It would allow an erroneous sequence number length field to be present without an erroneous sequence number field, which the specification indicates would not work, while requiring all erroneous sequence number fields to be associated with erroneous sequence number length fields, which the patent teaches is not necessary. The Board's construction, on the other hand, comports with what the patent teaches about the number and length fields. Even though the language of claim 15, standing alone, provides some support for Wi-Fi's interpretation, we hold that in the end the claim must be read as the Board construed it in order to be faithful to the invention disclosed in the specification.

Wi-Fi One, LLC v. Broadcom Corp., 887 F.3d 1329, 1346 (Fed. Cir. 2018).

A plain reading of the claim, in light of the specification, communicates that each device within the plurality generates a quality score or value for itself. Thus, if it is possible to have only one “other device[]” or “other electronic device[]” (per the above determination) then it is equally possible to have just one “quality score[]” or “second value[]” and still satisfy the claim.

Accordingly, “other devices” and “other electronic devices” in the 128 and 311 patents, respectively, do not mean “two or more,” and “quality scores” and “second values” in the 128 and 311 patents, respectively, do not mean “two or more.”

F. 128 Patent - “forgo outputting an audible response”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“refrain from outputting any and all audible responses”

JC at 2.

128 Patent - “forgoing outputting a response”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“refraining from outputting any and all responses”

JC at 2.

311 Patent - “[forgoing / forgo] responding”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	“[refrain]/[refraining] from outputting any and all responses”

JC at 2.

As with “command to temporarily refrain from performing speech recognition,” the parties dispute what sort of actions, if any, an electronic device may perform while it is “forgo[ing]” the responses listed in the three claim terms. Google rejects interpretations where the device “cannot perform any action whatsoever relating to a detected voice input, including ‘fulfilling’ a user request.” CIMB at 29. Google emphasizes that, pursuant to the specification, “the claimed ‘forgo[ing]’ is tied to a response to the detected voice input” with “[e]xamples of responding to a detected voice input include outputting an audible response (*e.g.*, ‘voice confirmation of the user request, reading out of the user-requested information’) and/or providing a visual response (*e.g.*, ‘a

notification or a LED light pattern’).” *Id.* at 30 (citing 128 patent at 21:61-65, 23:4-7, 29:59-30:12, 29:32-54). “In contrast,” Google continues, “a device that forgoes responding to a detected voice input may, for example, power down a speaker or associated amplifier.” *Id.* (citing 128 patent at 21:61-65, 23:4-7, 29:67-30:1). Google refers to this as fulfillment and contends it should not be considered a “response” contrary to Sonos’ position. *Id.* at 32; *see generally id.* at 30-33; CRMB at 11-13 (characterizing Sonos’ construction as forbidding “‘any and all’ *operation*” by the electronic device (emphasis added)).

Sonos views its use of “any and all” as properly clarifying that “excluding one type or instance of a response while permitting other types of instances of responses” does not fall within the claim. RIMB at 25. Sonos explains:

A simple example illustrates this point. Consider a device that outputs only a “beep” sound in response to the detected voice input. Under Google’s interpretation, it could be argued that, although the device output an audible response, it forwent outputting every other possible type or instance of audible responses, such as a “ding” audible response or even a second “buzz” audible response or a “buzz” audible response of a different duration, etc. Similarly, a device that outputs only a flashing light response could be argued to have forwent outputting a continuously-on light response or that it even forwent outputting a “buzz” or “ding” response. In this respect, no matter what response a device ultimately outputs, it can always be argued that it forwent outputting a different type of response. In other words, the device always forgoe outputting what it does not output.

Such an interpretation would therefore render the “forego”/“foregoing” phrases essentially meaningless or superfluous since it could be argued that these phrases are satisfied no matter what the device does.

Id. at 25-26 (citing *In re Power Integrations, Inc.*, 884 F.3d 1370, 1376 (Fed. Cir. 2018); *Bicon, Inc. v. Straumann Co.*, 441 F.3d 945, 950-51 (Fed. Cir. 2006); *Imaginal Systematic, LLC v. Leggett & Platt, Inc.*, 805 F.3d 1102, 1109 (Fed. Cir. 2015)). Sonos also finds support in the claims’ use of indefinite articles “a” and “an” before “response” to show they encompass “multiple types of responses rather than only excluding one particular type.” *Id.* at 26. Sonos also

rejects Google’s contention that fulfillment of a user request is different than responding to a user’s request based on the patents’ specification (RRMB at 13 (citing 311 patent at 7:48-63, 24:60-62, 17:6-7; CIMB at 31), 14 (“The claim recites ‘response’ and Sonos’s construction recites ‘response.’ . . . Exactly which responses need be forgone and which are allowable is apparently up to Google. This does not accord with proper claim interpretation.”)), and criticizes Google for “confusingly point[ing] to an unclaimed embodiment as apparent evidence that Sonos’s construction is incorrect for excluding this unclaimed embodiment.” *Id.* at 14 (citing CIMB at 33; 128 patent at 23:35-45).

First, it is a mistake to conflate and resolve at the same time these three terms which, on their faces, are markedly different. The very act which must be “forgone” is different between the 128 and 311 patents. In the 128 patent, the device must forgo “outputting” a response. 128 patent at cl. 1. Whereas in the 311 patent, the device must forgo “responding,” more generally. 311 patent at cl. 2. And then within the 128 patent, the first phrase very clearly forgoes outputting an “audible response” while the second forgoes outputting “a response,” again, more generally. 128 patent at cl. 1. Google pays no heed to these differences (*see generally* CIMB at 29-33; CRMB at 11-13), making much of its argument irrelevant.

Second, no party contends anything other than the plain and ordinary meaning controls, and the plain and ordinary meaning of the first term, “forgo outputting an audible response to the detected voice input,” is to refrain from or skip outputting a sound in response to the voice previously detected. Under this meaning, there is no restriction on the kind of sound—only that it have no connection to the detection of voice; *e.g.*, a previously scheduled clock chime may sound without falling outside of the limitation because it is not done in “response” to the voice input. Sonos’ proposed construction is not inconsistent with this ordinary meaning so long as it is

respected that “any and all audible responses” is limited to “responses” and not any sound whatsoever. It is also reasonable to understand “forgo” as “refrain” because one cannot forgo an act that was not otherwise going to occur (*e.g.*, one cannot “forgo” jumping twenty feet into the air), despite Google’s protestation. CIMB at 32 n.7.

Equally plain is that other kinds of responses to detected voice may be outputted and yet still fall within this term so long as those responses are not “audible”; *e.g.*, a user speaks and a red light flashes or a motor vibrates. Google’s concern that “a visual indication that a device is listening to a detected voice input, but not yet responding” or other “fulfillments” may get swept into this interpretation (CIMB at 33) is meritless for two reasons. One, the limitation only concerns “audible” responses—not light, motion, or other electrical communication. And two, a device which emits light to show that it is listening to voice is not necessarily a “response” to that voice—the light may have been already on because the device is listening all the time. On the other hand, a device which *turns on* a light to show that it heard a user’s voice would fall under the term, and rightly so.

Accordingly, there is nothing improper with Sonos’ construction of “refrain from outputting any and all audible responses” and it is accepted.

Turning to “forgoing outputting a response,” this obviously mirrors the previous limitation but omits the word “audible.” Thus, in a completely plain and ordinary sense, more than just audible responses are now blocked from being “output.” This would include turning on a light to show a device heard a user speak or even sending a “fulfillment” command to another device because of detected speech. There is no intrinsic evidence to suggest otherwise, and Google’s assertion that the patent differentiates between “responses” and “fulfillments” is unpersuasive. Its cited passages (CIMB at 32 (citing 128 patent at 9:38-43, 24:45-51, 10:2-20, 24:61-25:7)) do not

come close to the “exacting” standard set forth by *Thorner* and, more importantly, contradict Google’s position. The patent states that instructions sent between electronic devices in response to detected speech constitute a “response”:

For example, in response to a hotword voice input to awaken the electronic device 190, the electronic device 190 generates and presents (e.g., outputs) a voice greeting that includes the first user’s name (e.g., “Hello David,” “Good morning Mary”). The response may include a voice output *and/or instructions to another device to perform an operation*.

128 patent at 24:45-51 (emphasis added); *see id.* at 10:2-20 (disclosing that a message back to the user “that the request has been fulfilled” is a response).

Accordingly, there is nothing improper with Sonos’ construction of “refrain from outputting any and all responses” and it is accepted.

Lastly, “forgoing responding” is broader still. In a plain and ordinary sense, this means the device may not respond to the “detected input” at all, either by emitting an output or conducting other internal processes. In other words, the detected input must be ignored by that device. This does not mean the non-leader device must refrain from any activity whatsoever; *e.g.*, pre-scheduled clock chimes are permissible. It also does not mean responding to a leader device’s command to play a cat video, because while that is a form of response, it is a response to a signal other than the “detected input.” *See, e.g.*, CIMB at 31 (citing 128 patent at 9:38-43).

With that said, Sonos’ construction is not adopted entirely because it imports the word “outputting” from the other limitations when it was very clearly left out by the patentee. *Nystrom v. TREX Co., Inc.*, 424 F.3d 1136, 1143 (Fed. Cir. 2005) (“When different words or phrases are used in separate claims, a difference in meaning is presumed.”). Otherwise, the construction follows the familiar pattern and is accepted.

Accordingly, “[forgo / forgoing] responding” means “[refrain]/[refraining] from any and all responses” to the detected input.

G. 128 Patent - “in accordance with a determination that the quality score generated by the first electronic device is the highest . . . in accordance with a determination that the quality score generated by the first electronic device is not the highest”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	Indefinite

JC at 2.

311 Patent - “in accordance with a determination that the first value is higher than the second values . . . in accordance with a determination that the first value is not higher than the second scores”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	Indefinite

JC at 2.

Sonos identifies two pairs of claim limitations reciting similar subject matter—one pair in the 128 patent and one in the 311 patent—as indefinite under 35 U.S.C. § 112. These include:

in accordance with a determination that the quality score generated by the first electronic device is the highest amongst the generated quality score and received quality scores for the voice input

....

in accordance with a determination that the quality score generated by the first electronic device is not the highest amongst the quality scores for the voice input generated by the plurality of electronic devices, forgoing outputting a response to the detected voice input.

128 patent at cl. 1;

in accordance with a determination that the first value is higher than the second values, responding to the detected input.

....

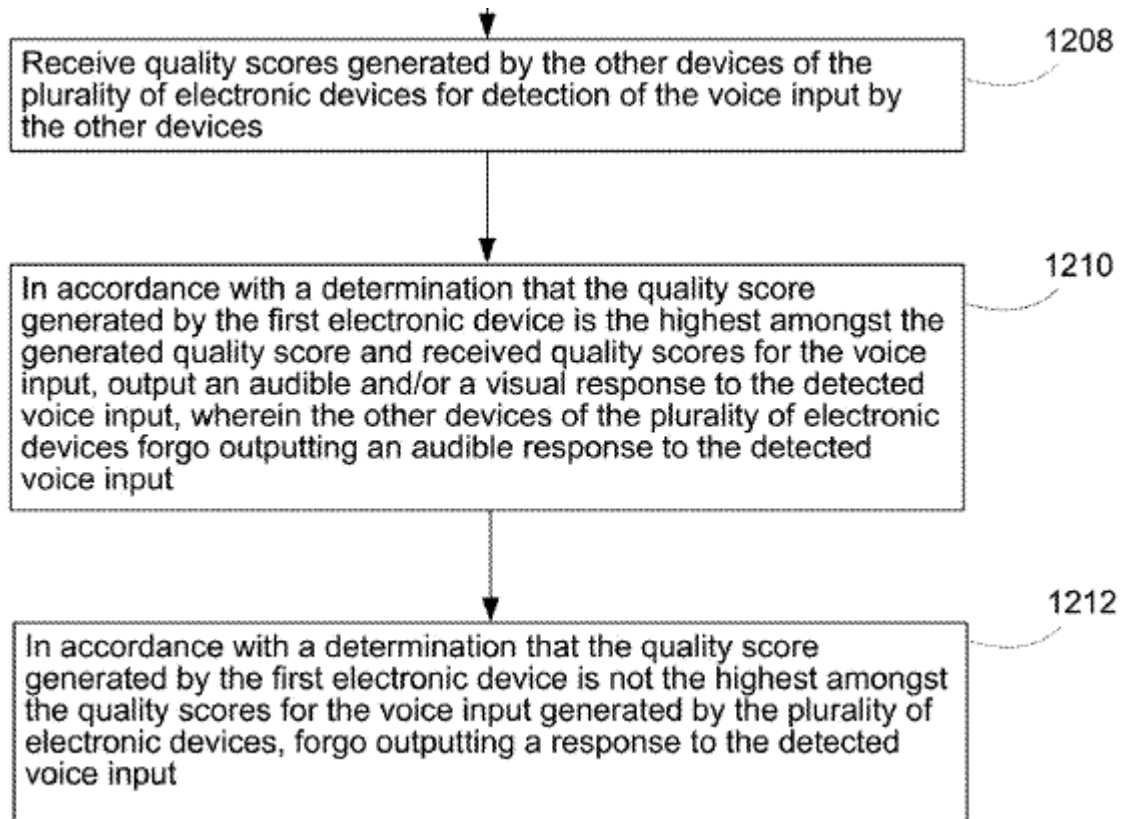
in accordance with a determination that the first value is not higher than the second scores, forgoing responding to the detected input.

311 patent at cls. 1, 2. Sonos contends, “[i]t is black letter law that method or process claims require the performance of each and ever recited step” and “[i]f, however, it is impossible to carry out each and every recited method step (*i.e.*, if two steps contradict each other), then that method claim is invalid for indefiniteness. . . .” RIMB at 31 (citations omitted), 34 (citing *Process Control Corp. v. HydReclaim Corp.*, 190 F.3d 1350, 1359 (Fed. Cir. 1999)). Sonos argues it is impossible to carry out the method of claim 1 of the 128 patent and claim 2 of the 311 patent because the limitations above are “alternative and contradictory steps”—the first clause recites a score/value which is the highest and the second clause recites that same score/value as not the highest. *See id.* at 32-33.

Sonos rejects viewing the limitations as otherwise permissible conditional method steps, as argued by Google (CIMB at 34), because they “fail to use any hallmark of conditional language, such as ‘if,’ ‘when,’ or ‘on a condition’” RRMB at 15 (citing RRMB, Ex. M at ¶ 32-33); *see* RRMB, Ex. M at ¶ 35 (citing 128 patent at Fig. 6)). If they are so viewed, Sonos argues it would constitute an improper rewriting of the claims and leave claim 2 of the 311 patent indefinite for broadening the scope of its independent claim. RRMB at 16-17 (citing 35 U.S.C. § 112(d); *Multilayer Stretch Cling Film Holdings, Inc. v. Berry Plastics Corp.*, 831 F.3d 1350, 1362 (Fed. Cir. 2016); *Pfizer, Inc. v. Ranbaxy Labs. Ltd.*, 457 F.3d 1283, 1291 (Fed. Cir. 2006)).

Sonos’ contentions are rejected. Sonos does not dispute that conditional method steps are possible, it just disputes that the intrinsic evidence supports reading the “in accordance with a determination . . .” limitations in this way. To the contrary, the shared specification of the two patents is clear that for any given voice input, all detecting devices within the multi-device network evaluate whether they have the highest “score” or “value” as compared to the other

members of the group, and if so, they are the only device which will respond or further process the voice input. *See, e.g.*, 128 patent at 22:66-23:45 (describing “leadership determination process” and “just the leader device 1002-1 outputs the response per the leadership determination”); 311 patent at 23:9-55 (same). This is true even for Figure 6, cited by Sonos (RIMB at 33), which on its face shows both “determination” method steps somehow occurring in a single process flow:



128 patent at Fig. 6; 311 patent at Fig. 6. The patent nonetheless teaches that, in the context of this figure, each device compares its “score” or “value” to those obtained from the other devices, so that only one device responds:

In either situation, the device with the highest score (e.g., device 1002-1) outputs the response, and the other devices (e.g., devices 1002-2 and 1002-3), which have the lower scores, do not output the response. In some implementations, the electronic device that is the leader (e.g., the device with the highest score in the example here) also is the leader with respect to processing or pre-processing subsequent voice inputs from the user.

128 patent at 29:50-58; 311 patent at 29:62-30:3. Sonos points to no embodiments where more than one leader is established (*i.e.*, the “determination” limitations are not conditional alternatives). *See generally* RIMB at 29-34; RRMB at 15-18. In fact, it acknowledges there can be only one, stating “[i]ndeed, the point of the invention is to have *multiple* devices carry out the leader election process in order to select one of them as the leader.” RRMB at 19 (emphasis in original). This is sufficient to dispense with Sonos’ primary argument.

Additionally, Google points out that the patentee effectively acts as its own lexicographer for the phrase “in accordance with a determination that.” CIMB at 35. In the conclusion of the specification, the 128 patent equates the basic conditional signal “if” with the claimed phrase “in accordance with a determination that”:

As used herein, the term “if” is, optionally, construed to mean “when” or “upon” or “in response to determining” or “in response to detecting” or “in accordance with a determination that,” depending on the context. Similarly, the phrase “if it is determined” or “if [a stated condition or event] is detected” is, optionally, construed to mean “upon determining” or “in response to determining” or “upon detecting [the stated condition or event]” or “in response to detecting [the stated condition or event]” or “in accordance with a determination that [a stated condition or event] is detected,” depending on the context.

128 patent at 35:6-16; *see* 311 patent at 35:18-28. Sonos’ only response, given in a footnote at the end of its rebuttal brief, is that a POSITA would not understand the claims as implementing an “if” because the paragraph states “if” should be construed as “in accordance with a determination that,” and not the other way around. RRMB at 18 n.10 (citing RRMB, Ex. M at ¶¶ 38-41)). Sonos’ expert also highlights the paragraph’s use of “optionally” and “depending on the context” as additional reasons why a POSITA would not view the challenged limitations as “if” statements. RRMB, Ex. M at ¶¶ 38-41.

Sonos’ objections are not persuasive. While not explicit, the spirit of the paragraph is that “if” and “in accordance with a determination that” are variations on the same theme and therefore

interchangeable. And it is reasonable to find that a POSITA would consider this substitution when encountering the distinctive “in accordance with a determination that” phrase within the claims. When added to the other intrinsic evidence discussed above, it is clear the two claim limitations are intended to be and would be understood as alternative method steps.

Sonos’ additional argument—that treating the two method steps as alternatives creates improper broadening problems with claim 2 in the 311 patent—is equally unavailing. The scope of claim 1 does not grow when the language of claim 2 is introduced. As it stands, claim 1 covers the (abridged) steps of: detecting a voice input; determining a first value; receiving second values; and “in accordance with a determination that the first value is higher than the second values, responding to the detected input.” 311 patent at cl. 1. Taking just this claim into consideration, if a device performs the first three steps and the recited condition (“first value” is higher than all other second values) does not occur, the device is free to do anything else. When claim 2 is added, that freedom disappears. When the “first value” is not higher, the device must “forgo[] responding to the detected input.” Thus, the scope of the combined claim 1 and 2 reads on a more limited set of actions and is, by definition, narrower.

Accordingly, the “in accordance with a determination” limitations are not indefinite for reciting contradictory steps.

H. 128 Patent - “cause the electronic device”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	Indefinite

JC at 3.

Sonos alleges the phrase “cause the electronic device” appearing in claims 11 and 15 of the 128 patent is indefinite for lack of antecedent basis. RIMB at 34-35. Sonos explains:

Here, there is no implication as to what specific “electronic device(s)” the recitation “the electronic device” refers. In fact, there are two equally plausible interpretations on how a POSITA might interpret this claim. Ex. F (White Dec.), ¶¶ 34-39. When this is the case, the Court must hold the claim indefinite because the claims do not reasonably inform a POSITA as to the scope of the claim. *Honeywell Intern. Inc. v. ICM Controls Corp.*, 2014 WL 4248434, *8-*10 (D. Minn. 2014) (claim indefinite where it was clear that some language was missing from the preamble, and it was not clear how to correct the missing language where there were two plausible choices); *see also Novo Indus., L.P. v. Micro Molds Corp.*, 350 F.3d 1348, 1357 (Fed. Cir. 2003) (“A district court can correct a patent only if (1) the correction is not subject to reasonable debate based on consideration of the claim language and the specification and (2) the prosecution history does not suggest a different interpretation of the claims.”).

The claim is either meant to recite (i) “cause the [first] electronic device to perform operations comprising . . .” or “cause the electronic device[s] to perform operations comprising . . .” Either of these corrections would alleviate the antecedent issue – but because both are equally plausible, the claim cannot be corrected or interpreted and must be held indefinite.

Id. at 35; RRMB at 19 (“There is no indication one way or the other as to whether ‘the electronic device’ was intended to refer to ‘the [first] electronic device’ or ‘the electronic device[s].’” (emphasis removed)).

Sonos is incorrect. A plain reading of the term in context reveals “the electronic device” refers to the “first electronic device.” The preamble of claim 11, the first appearance of the term, reads:

A non-transitory computer readable storage medium storing one or more programs, the one or more programs comprising instructions, which, when executed by a first electronic device of a plurality of electronic devices, each of the plurality of electronic devices comprising one or more microphones, a speaker, and one or more processors, cause the electronic device to perform operations comprising:

128 patent at cl. 1. By its own language, this is a computer memory claim with instructions that are “executed by a first electronic device.” It is natural, therefore, to read the subsequent phrase “cause the electronic device to perform operations comprising,” as referring back to that first electronic device. The “plurality of electronic devices,” on the other hand, are not recited as

carrying out those instructions or executing them. Rather, they are merely the environment in which the first electronic device (with the claimed computer readable storage medium) operates, that is, it is one of a plurality of presumably similar devices that share quality scores with each other. Claim 15 solidifies this point, as there is no recitation of that plurality-device environment at all. 128 patent at cl. 15 (“The computer readable storage medium of claim 11, further comprising instructions which, when executed by the electronic device, cause the electronic device to perform operations comprising:”). Contrary to Sonos’ suggestions, there are no fatal drafting errors here (RRMB at 15) and the interpretation involves no rewriting of the claim (RIMB at 37).

Accordingly, “cause the electronic device” is not indefinite for lack of antecedent basis.

I. 311 Patent - “the second scores”

Google’s Construction	Sonos’ Construction
“the second values”	Indefinite

JC at 3.

As with “cause the electronic device,” Sonos contends “the second scores” in claim 2 of the 311 patent is indefinite for lack of antecedent basis. RIMB at 37 (citing *Halliburton Energy Servs., Inc.*, 514 F.3d at 1249). Sonos explains, “Claim 2 of the ‘311 patent recites ‘the second scores,’ but there is no reference to ‘scores’ in either claim 2 or claim 1 (from which claim 2 depends). Indeed, claim 1 references ‘second values,’ but the Applicant chose to claim ‘second scores.’” *Id.*

Google acknowledges the discrepancy but wants the term construed as “the second values” to correct “an obvious clerical error.” CIMB at 40. Google continues:

Sonos’s only stated basis for contending this term is indefinite is lack of antecedent basis. *See* Ex. 10 (Sonos Invalidity Contentions) at 7. Sonos advocates for the draconian result of indefiniteness in the face of a record that shows the Applicant’s intent—and inadvertent failure—to correct “the second scores” to “the second values” during prosecution. But Sonos’s argument should be rejected

here, as the error is obvious and a POSITA would understand the scope of the claim with reasonable certainty. Ex. 6 (Anderson Decl.) at ¶ 131. Sonos’s arguments should be rejected.

Id.; CRMB at 18.

Google and Sonos are both correct. There is an obvious error in claim 2 of the 311 patent. So obvious in fact, Google should have applied for a certificate of correction shortly after the patent issued or, at the very least, in preparation for the filing of its complaint. This is because, unlike a District Court, the Commission lacks the statutory authority to correct even obvious typographical errors. *See Novo Indus.*, 350 F.3d at 1357 (district courts may “correct obvious minor typographical and clerical errors” in patents, but not “major errors”); *see, e.g., Wind Turbines*, Inv. No. 337-TA-641, Comm’n Op. at 36 (“The Commission lacks the authority to correct inventorship under Section 256 or any other statutory provision, and the Commission’s authority in this regard must be conferred by statute.”).

Accordingly, “the second scores” is indefinite for lack of antecedent basis.

J. 311 Patent - “the detected output”

Google’s Construction	Sonos’ Construction
“the detected input”	Indefinite

JC at 3.

Again, Sonos argues “the detected output” is indefinite for lack of antecedent basis. Sonos explains, “Claim 12 recites ‘communicating to the other devices of the plurality of electronic devices to forgo responding to the detected output.’ However, there is no reference to a ‘detected output’ anywhere in claim 12 or in claim 10 (from which claim 12 depends).” RIMB at 39; RRMB at 20-21. Google again paints this as “an obvious typographical error” which should be corrected to “the only reasonable interpretation . . . ‘the detected input.’” CIMB at 42; *see* CRMB

at 19-20. For the same reasons provided in connection with “the second scores,” correction of claim language mistakes—obvious or otherwise—is not within the Commission’s power. *See Novo Indus.*, 350 F.3d at 1357; *Wind Turbines*, Inv. No. 337-TA-641, Comm’n Op. at 36.

Accordingly, “the detected output” is indefinite for lack of antecedent basis.

K. 615 Patent - “short-range signal transmission path”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	Indefinite

JC at 4.

Sonos contends “short-range signal transmission path” within the 615 patent claims is indefinite. RIMB at 41. Sonos first questions whether “short-range” applies to the “signal” or the “path,” but reasons that in either interpretation it is indefinite “because neither the specification nor the file history gives ‘objective boundaries’ that would allow a person of skill in the art to discern what qualifies as ‘short.’” *Id.* Sonos acknowledges the specification’s listing of various technologies that could fill the role (*id.* (citing 615 patent at 8:42-67)) but faults the intrinsic evidence for not “provid[ing] any standard to determine when a signal might or might not qualify. Nor does it delineate the negative boundary by providing examples of signaling technologies that do *not* qualify as ‘short-range’” (*id.* at 42; *see id.* at 42-43 (citing RIMB, Ex. F at ¶ 96-100); *see* RRMB at 21, 22-24 (discussing failures of “proximity-based” definition)). Similarly, Sonos acknowledges that “academics often use the term ‘short-range’” but argues “they don’t use it in a single, consistent way that provides the kind of ‘reasonable certainty’ that *Nautilus* requires.” RIMB at 43 (citing RIMB, Ex. F at ¶ 93). To further illustrate the confusion, Sonos presents allegedly contradictory positions taken by Google in this investigation and in an unrelated district court case as to whether Wi-Fi qualifies as “short-range.” RRMB at 22 (citing *Uniloc 2017 LLC*

v. Google LLC, No. 2:18-CV-00493-JRG-RSP, 2020 U.S. Dist. LEXIS 9748 (E.D. Tex. Jan. 20, 2020)).

Google contends the term is not indefinite. It first points to the purpose of the path, which is to link “a ‘wireless device’ that is part of a communication network, and a ‘target device’ that is not yet part of that communication network.” CIMB at 45. Google then highlights those same protocol examples from the specification and argues they would enable a POSITA to understand the term. *Id.* at 46 (citing 615 patent at 8:9-67). According to Google, “[i]n view of these disclosures, the ’615 patent more than provides objective benchmarks that allows ‘a skilled artisan to compare’ a product’ with the examples in the specification to determine’ if a transmission path is a short-range transmission path.” *Id.* at 47 (citing *Sonix Tech. Co. v. Publications Int’l Ltd.*, 84 F.3d 1370, 1379 (Fed. Cir. 2017); *Niazi Licensing Corp. v. St. Jude Med. S.C., Inc.*, 30 F.4th 1339, 1349 (Fed. Cir. 2022)); CRMB at 20-21 (citing CRMB, Mahon Decl. at ¶ 52). Google further points to dependent claim 5, explicitly reciting many of the same “short-range” communication examples, and reasons that this proves claim 5 cannot be indefinite even if claim 1 is so held. CRMB at 23 (citing *Default Proof Credit Card Sys., Inc. v. Home Depot U.S.A., Inc. (d/b/a The Home Depot)*, 412 F.3d 1291, 1297 (Fed. Cir. 2005)).

Google also presents extrinsic evidence to show “‘short-range’ is [] common in the field and well understood by persons of skill,” including patent applications from Sonos and from Sonos’ testifying expert which use the term. CIMB at 46-47, 47 n.19 (citing CIMB, Exs. 13, 14)); CRMB at 21-22 (citing CRMB, Ex. 23). Google avers, “[t]his dooms Sonos’s indefiniteness allegations.” CRMB at 22.

Indefiniteness, and all other grounds of patent invalidity, must be proved by clear and convincing evidence. *Microsoft Corp. v. i4i Ltd. P’ship*, 564 U.S. 91, 95 (2011). Sonos has not

met this standard, because it is not clear or convincing that a POSITA would not be able to discern reasonable, objective, bounds to “short-range signal transmission path.” As for intrinsic evidence, it is notable that the claims use the term “short-range” but the 651 patent specification actually does not. *See generally* 615 patent. Even so, the parties’ briefing demonstrates there is no dispute that the specification’s list of communication technologies/protocols correspond to the claim element and thus qualify as “short-range.” *See* CIMB at 46; RIMB at 42. These examples, also appearing in dependent claim 5, contribute to informing a POSITA on the scope of the term. *See Niazi*, 30 F.4th at 1348 (“We explained that, while the claims did not define ‘not interfering substantially’ with mathematical precision, the intrinsic evidence provided guideposts for a skilled artisan to determine the scope of the claims.”) (citing *Enzo Biochem, Inc. v. Applera Corp.*, 599 F.3d 1325, 1335 (Fed Cir. 2010)).

The weight of the parties’ extrinsic evidence also favors Google. Sonos provides a declaration from its expert, Dr. Jules White, who, in relevant part, states “although academic literature sometimes uses the term ‘short-range’ when describing wireless technology, that term is always used in a particular context and does not have a consistent definition.” RIMB, Ex. F at ¶ 93. One of Dr. White’s cited exhibits, however, undercuts the claim. That exhibit is a 2009 textbook entitled *Short-Range Wireless Communications: Emerging Technologies and Applications*. RIMB, Ex. F, Exhibit 3. The introduction explains that short-range applies to a variety of contexts but in all cases shares certain characteristics including “wireless connectivity within a local sphere of interaction,” and is thus contrasted with cellular systems:

Short-range communications systems characterize a wide range of scenarios, technologies and requirements. There is no formal definition of such systems though one can always classify short-range systems according to their typical reach or coverage. We define shortrange communications as the systems providing wireless connectivity within a local sphere of interaction. Such a space corresponds to the first three levels of the multisphere model as discussed in the

Book of Visions of the Wireless World Research Forum (WWRF) [1]. Figure 1.1 depicts the multisphere concept, highlighting the levels associated with shortrange communications, namely Personal Area Network, Immediate Environment and Instant Partners [2–4]. Short-range systems involve transfer of information from millimeters to a few hundreds of meters. However, short-range communication systems are not only systems providing wireless connectivity in the immediate proximity, but in a broader perspective they also define technologies used to build service access in local areas. The WWRF envisions that by year 2017 there would be seven trillion wireless devices serving seven billion people. Certainly, the overwhelming majority of these devices will be short-range communication systems providing wireless connectivity to humans and machines.

Together with wide/metropolitan area cellular systems, short-range systems represent the two main developing directions in today’s wireless communications scene. In terms of design rules and target capabilities, short-range systems have certain commonalities as well as marked differences from their counterparts, cellular systems.

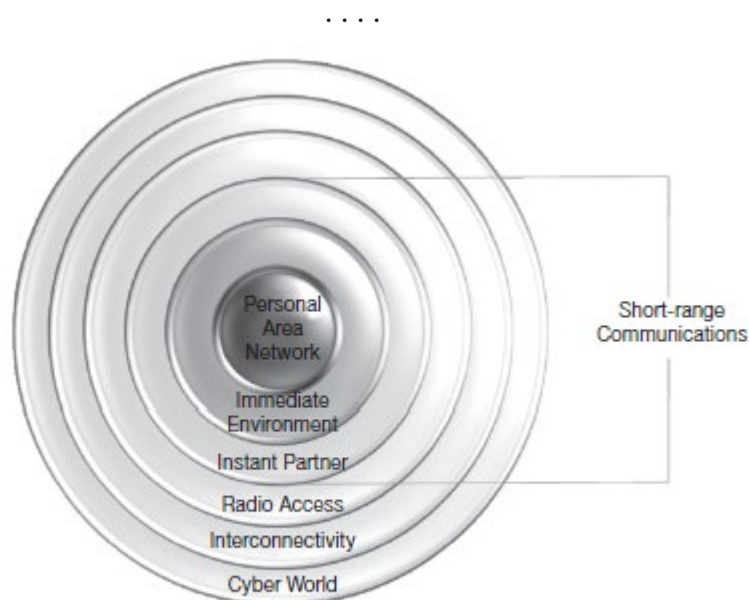


Figure 1.1 Short-range communications within the WWRF multisphere reference model.

RIMB, Ex. F, Exhibit 3 at 1-2. Admittedly, this textbook teaches that “short-range” is a broad concept, but broad does not mean indefinite. *Niazi*, 30 F.4th at 1347. Added to this is Google’s own collection of extrinsic evidence, including the opinions of its expert in the field, teaching that “short-range” is commonly used and understood to refer to local, low interference-causing communications. *See, e.g.*, CRMB, Ex. 17 at ¶¶ 39, 47, 51-58 (documentary citations omitted).

Overall, a clear and convincing case has not been made that a POSITA would not understand with reasonable certainty what “short-range signal transmission path” means. The evidence shows it is certainly not a term coined by the 615 patent.

Accordingly, “short-range signal transmission path” is not indefinite.

L. 615 Patent - “in response to receiving the request, receiving one or more network configuration packets”

Google’s Construction	Sonos’ Construction
No construction necessary (plain and ordinary meaning)	Indefinite

JC at 4.

Sonos contends “in response to receiving the request, receiving one or more network configuration packets” is indefinite because “it is not at all clear how a target device could be said to receive network configuration packets as a result of receiving a request to start the commissioning process.” RIMB at 46. Sonos continues:

The claims provide no indication of what the target device must do “in response to” receiving the request that results in it receiving the network configuration packets. And the only possibility—establishing the short-range communication path—is already expressly recited in the claims. Consequently, the disputed term is indefinite. *See USB Bridge Sols., LLC v. Buffalo Inc.*, No. 1-17-cv-0011558, 2020 U.S. Dist. LEXIS 67678, at *20 (W.D. Tex. Apr. 17, 2020) (determining “conversion logic” is indefinite because the term is “never explained to perform [its] function”; rather “other components perform” the claimed conversion function).

Id. at 48. Sonos rejects an interpretation that equates “in response” to “after,” as this would improperly read out the causality invoked by the term. *Id.* (“‘In response to’ indicates a different relationship that goes beyond temporal sequencing.”). Finally, Sonos argues the intrinsic evidence provides no answers here as “the specification includes no description—even using different words—of what is meant by receiving network configuration packets in response to receiving the request.” *Id.* at 49 (citing RIMB, Ex. F at ¶¶ 108-109).

Sonos is not persuasive. If its chief complaint is that “[t]he claims provide no indication of what the target device must do ‘in response to’ receiving the request that results in it receiving the network configuration packets,” then the answer is straightforward—anything. The claim leaves unstated what step or steps occur in between “receiving the request” and “receiving one or more network configuration packets.” This does not make for indefiniteness; it makes for a broad term. CRMB at 24 n.16 (citing *Solvay S.A. v. Honeywell Int’l Inc.*, 742 F.3d 998, 1005 (Fed. Cir. 2014) (“The well-established meaning of ‘comprising’ in a method claim indicates that the claim is open-ended and allows for additional steps.”)).

Alternatively, if Sonos’ complaint is that the patent fails to discuss “receiving one or more network configuration packets” in response to “receiving the request” (*see* RIMB at 49), then that is a written description problem, not an indefiniteness problem. And if the complaint is that a POSITA would not know how to achieve “receiving one or more network configuration packets” in response to “receiving the request” (*see id.* at 46) without undue experimentation, then that is an enablement problem—again, not indefiniteness. Google puts it succinctly and accurately: “[t]he claim means what it says: network configuration packets are received as a result of, and because of, the request to initiate commissioning.” CRMB at 25.

Further, the term is unlike that found indefinite in *Trusted Knight Corp v. Int’l Bus. Machines Corp.*, No. 14-1063-LPS-CJB, 2015 WL 7307134 (D. Del. Nov. 19, 2015), even though both use the phrase “in response to.” *See* RIMB at 47-48. In that case, the “in response to” came at the very end of a computer memory claim comprising certain software processes:

[S]aid software processes including:

a process of detecting a browser form . . .

a process of intercepting data inputs keyed in by a user . . .

a process of (1) submitting the keyed-in data to a designated entity through the API stack while (2) clearing confidential data from intercepted data . . . in response to the software key logging through the API stack to an internet communication port.

U.S. Patent No. 8,316,445 at cl. 1; see *Trusted Knight*, 2015 WL 7307134 at *4-6. The phrase was determined to be indefinite because it was unclear grammatically about “what *is* happening ‘*in response to* the software key logging?’” *Id.* at *5 (emphasis in original). Here, on the other hand, the answer is straightforward. A second reception happens in response to the first reception via unclaimed means. RIMB at 48. Again, a broad claim should not be mistaken for an indefinite claim. *Niazi*, 30 F.4th at 1347.

Accordingly, “in response to receiving the request, receiving one or more network configuration packets” is not indefinite.

SO ORDERED.



Cameron Elliot
Administrative Law Judge