UNITED STATES INTERNATIONAL TRADE COMMISSION
WASHINGTON, D.C.

Before the Honorable Cameron Elliot
Administrative Law Judge

IN THE MATTER OF CERTAIN AUDIO PLAYERS AND COMPONENTS THEREOF II	**Investigation No. 337-TA-1330**

COMPLAINANT'S INITIAL MARKMAN BRIEF

SONOS EXHIBIT 1010

TABLE OF CONTENTS

TABLE OF CONTENTS ii
TABLE OF AUTHORITIES iv
TABLE OF EXHIBITS vi
I. INTRODUCTION 1
II. TECHNOLOGY BACKGROUND 2
A. Voice Assistant Technologies Relating to the '748, '128, and '311 Patents 3
B. Device Commissioning Technologies Relating to the '615 Patent 7
III. LEVEL OF ORDINARY SKILL IN THE ART 8
IV. LEGAL STANDARDS 9
A. Claim Construction 9
B. Indefiniteness 10
C. Correcting Obvious Errors 10
V. '748 Patent 10
A. Overview of the '748 Patent 10
B. Agreed Constructions 11
C. "[perform/performing] speech recognition" (All Claims) 12
D. "a command to temporarily refrain from performing speech recognition" (All Claims) 15
E. "predetermined spoken audio command" (All Claims) 17
F. "[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" (All Claims) 19
VI. '128 and '311 Patents 23
A. Overview of the '128 and '311 Patents 23
B. "other devices of the plurality of electronic devices" / "quality scores generated by the other devices of the plurality of electronic devices" ('128 All Claims) 25
C. "second values … from other electronic devices" / "second values from other electronic devices of the plurality of electronic devices" ('311 All Claims) 28
D. "forgo outputting an audible response" / "forgoing outputting a response" ('128 All Claims) / "[forgoing / forgo] responding" ('311 Claims 2, 3, 12) 29
E. "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" ('128 Claims 1, 2, 3, 5) / "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" ('311 Claim 2) 33
F. "cause the electronic device" ('128 Claims 11, 12, 13, 15) 37
G. "the second scores" ('311 Claims 2, 3) 39
H. "the detected output" ('311 Claim 12) 42
VII. '615 Patent 44
A. Overview of the '615 Patent 44
B. "short range signal transmission path" (All Claims) 45
C. "in response to receiving the request, receiving one or more network configuration packets" (All Claims) 47
VIII. CONCLUSION 49

TABLE OF AUTHORITIES

Page

Cases

AIA Eng'g Ltd. v. Magotteaux Int'l S/A,
657 F.3d 1264 (Fed. Cir. 2011) 36

Bicon, Inc. v. Straumann Co.,
441 F.3d 945 (Fed. Cir. 2006) 21

Certain Multiple Mode Outdoor Grills and Parts Thereof, Inv. No. 337-TA-895,
Comm'n Op., 2015 WL 13817118 (Feb. 3, 2015) 10

Cheese Sys., Inc. v. Tetra Pak Cheese & Powder Sys., Inc.,
725 F.3d 1341 (Fed. Cir. 2013) 26

Cybersettle, Inc. v. Nat'l Arbitration Forum, Inc.,
243 F. App'x 603 (Fed. Cir. 2007) 34

Dayco Prods., Inc. v. Total Containment, Inc.,
258 F.3d 1317 (Fed. Cir. 2001) 27

Eidos Display, LLC v. AU Optronics Corp.,
779 F.3d 1360 (Fed. Cir. 2015) 39

Festo Corp. v. Shoketsu Kinzoku Kogyokabushiki Co.,
535 U.S. 722 (2002) 10

GE Lighting Sols., LLC v. AgiLight, Inc.,
750 F.3d 1304 (Fed. Cir. 2014) 28

Helmsderfer v. Bobrick Washroom Equip., Inc.,
527 F.3d 1379 (Fed. Cir. 2008) 9

Hoffer v. Microsoft Corp.,
405 F.3d 1326 (Fed. Cir. 2005) 10

Howmedica Osteonics Corp. v. DePuy Orthopaedics, Inc.,
No. 11-CV-6498 SDW MCA, 2013 WL 3455727 (D.N.J. July 9, 2013) 44

Hytera Commc'ns Co. v. Motorola Sols., Inc.,
841 F. App'x 210 (Fed. Cir. 2021) 34, 37

Interactive Gift Express, Inc. v. Compuserve, Inc.,
256 F.3d 1323 (Fed. Cir. 2001) 9

Interdigital Tech. Corp. v. Lenovo Holding Co., Inc.,
No. CV 19-1590-LPS, 2021 WL 1856937 (D. Del. May 10, 2021) 35, 37

Joy MM Delaware Inc. v. Cincinnati Mine Mach. Co.,
No. CIV.A. 09-1415, 2010 WL 3984822 (W.D. Pa. Oct. 9, 2010), *aff'd*, 497 F. App'x
970 (Fed. Cir. 2012) 38

K-2 Corp. v. Salomon S.A.,
191 F.3d 1356 (Fed. Cir. 1999) 9, 20

Lincoln Nat. Life Ins. Co. v. Transamerica Life Ins. Co.,
609 F.3d 1364 (Fed. Cir. 2010) 37

Merck & Co. v. Teva Pharms. USA, Inc.,
395 F.3d 1364 (Fed. Cir. 2005) 21

Microsoft Corp. v. I4I Limited Partnership,
564 U.S. 91 (2011) 10

Nautilus, Inc. v. Biosig Instruments, Inc.,
572 U.S. 898 (2014) 10

Niazi Licensing Corp. v. St. Jude Med. S.C., Inc.,
30 F.4th 1339 (Fed. Cir. 2022) 47

Oatey Co. v. IPS Corp.,
No. 07-1214 (Fed. Cir. Jan. 30, 2008) 29

Omega Eng'g, Inc, v. Raytek Corp.,
334 F.3d 1314 (Fed. Cir. 2003) 22

Omeprazole Patent Litigation,
84 F. App'x 76 (Fed. Cir. 2003) 27

Phillips v. AWH Corp.,
415 F.3d 1303 (Fed. Cir. 2005) 9

Sonix Tech. Co. v. Publications Int'l, Ltd.,
844 F.3d 1370 (Fed. Cir. 2017) 47

Thorner v. Sony Computer Entm't Am. LLC,
669 F.3d 1362 (Fed. Cir. 2012) 9

TQ Delta, LLC v. 2Wire, Inc.,
373 F. Supp. 3d 509 (D. Del. 2019) 49

U.S. Surgical Corp. v. Ethicon, Inc.,
103 F.3d 1554 (Fed. Cir. 1997) 9, 17

Versa Corp. v. Ag-Bag Int'l Ltd.,
392 F.3d 1325 (Fed. Cir. 2004) 26

York Prods., Inc. v. Cent. Tractor Farm & Family Ctr.,
99 F.3d 1568 (Fed. Cir. 1996) 26

TABLE OF EXHIBITS

Exhibit No.	Description
1	U.S. Patent No. 9,632,748
2	U.S. Patent No. 9,812,128
3	U.S. Patent No. 11,024,311
4	U.S. Patent No. 11,050,615
5	*The Smart Audio Report*, NPR (2020)
6	Declaration of David V. Anderson
7	Fernández et al., *An Application of Recurrent Neural Networks to Discriminative Keyword Spotting*, ICANN 2007
8	U.S. Patent No. 9,318,107
9	Morgan and Bourlard, *Continuous Speech Recognition*, IEEE Signal Processing Magazine (May 1995)
10	Attachment A (Excerpts) to Sonos's December 7 Supplemental Responses to Google's First Interrogatories (Invalidity Contentions)
11	Keshet, *Automatic Speech and Speaker Recognition* (2009)
12	File History (Excerpts) of U.S. Patent No. 11,024,311
13	U.S. Patent App. Pub. No. 2022/0244905
14	U.S. Patent App. Pub. No. 2022/0224971

Pursuant to Order No. 6, Complainant Google LLC ("Complainant" or "Google") respectfully submits this Initial Markman Brief for disputed terms in the claims of U.S. Patent Nos. 9,632,748 ("the '748 patent"), 11,024,311 ("the '311 patent"), 9,812,128 ("the '128 patent"), and 11,050,615 ("the '615 patent") (collectively, the "Asserted Patents") (Exhibits 1-4, respectively).

I. INTRODUCTION

Across the Asserted Patents, the meaning of all but two of the thirteen terms proposed by Respondent Sonos, Inc. ("Sonos") is so plain that the terms do not require construction. (Two require simple corrections as the result of obvious drafting errors.) Sonos's approach takes one of two forms: Sonos either contorts the claims with lengthy claim construction proposals that—when stripped down to their core—vitiate claim limitations, inject new limitations unsupported by the intrinsic evidence, or both; or, seizing on contrived inconsistencies or clerical errors that a person of ordinary skill in the art would easily have been able to understand, Sonos argues that the claims are indefinite.

For instance, with respect to the '748 patent, Sonos's proposal for "[determine/determining]…based at least in part on information…" omits the "based at least in part on" clause altogether, and adds a new negative limitation ("without any of the plurality of computing devices processing…") that has no support in the intrinsic evidence. For the '128 and '311 patents, Sonos asserts that a POSITA would be unable to understand method claims reciting alternative conditional steps, when the specifications describe precisely the claimed methods using the exact same language. Similarly, for the '615 patent, Sonos ignores the plain claim language and disclosures in the specification to assert that the term "short range signal transmission path" is indefinite.

For the reasons set forth herein, Google respectfully requests that the ALJ adopt its proposals, and decline to adopt Sonos's constructions.

II. TECHNOLOGY BACKGROUND

In recent years, voice assistants—software that generally allows interaction with, and control of, electronic devices by audible phrases and commands—have exploded in popularity. Ex. 5 (*The Smart Audio Report*, NPR) at 5-6, 16. Google has been a pioneer in the development and advancement of voice-assistant technology since the early 2010s. Google's innovations began with software that enabled voice-based search on Android smartphones and the Chrome web browser, and evolved into Google Assistant, a service that gives users a convenient interface for performing tasks on smart phones, speakers, watches, televisions, and more. Google Assistant, which is part of the domestic industry in this Investigation, has transformed smart device management, connected homes, and mobile computing.

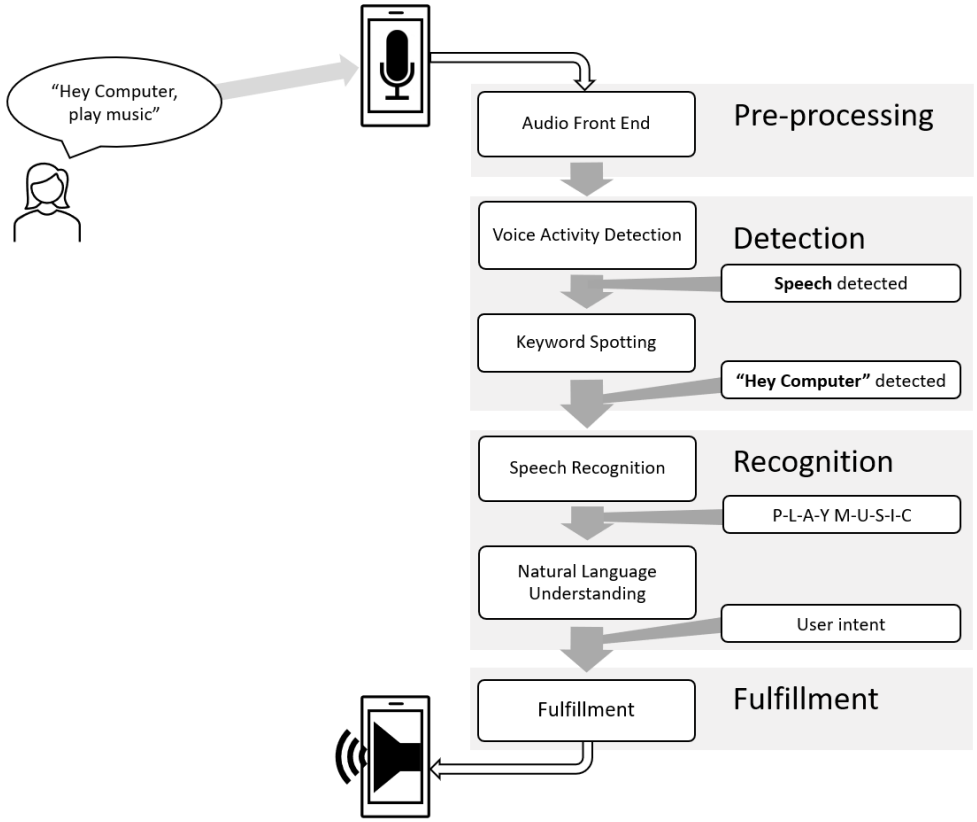
As voice assistant technology was implemented in a greater variety of devices, problems arose from the use of the technology in multiple devices in the same environment. For example, an open living space may have more than one device running a voice assistant application—e.g., a smart television and smart speaker, along with a smartphone and/or watch. As adoption of technology is and was often premised on its ease of use, and the number of smart devices (including those supporting voice assistant technology) began to increase, it became more important for new devices to be easily introduced (or "commissioned") into an environment (e.g., a home network). It was also important that these newly-introduced devices (or multiple devices in a particular environment) did not interfere with one another in the execution of their respective voice assistant applications. If commissioning a new device onto a network was difficult, or if multiple devices' voice assistant applications interfered with one another, the user's experience would be adversely affected. The innovations of the Asserted Patents are designed to enhance these operations.

Below is a background summary of technologies relevant to the Asserted Patents and claim construction disputes.

A. Voice Assistant Technologies Relating to the '748, '128, and '311 Patents

Voice assistants and the smart devices that support them process audio in multiple stages—a "pipeline"—that generally increase in computational intensity and processor usage as the pipeline proceeds. A voice assistant pipeline typically starts with a device's microphones receiving a user request and concludes with fulfillment of the request (i.e., execution of a requested action) by (a) the device that received it or (b) another device at receiving device's direction. As shown below, the stages of the pipeline can be grouped together by function (with the understanding that not every "pipeline" in a given system necessarily includes every stage). Ex. 6 (Declaration of David V. Anderson (hereafter, "Anderson Decl.")) at ¶¶ 36-37.

In general, the pipeline will include a pre-processing function ("audio front end"), detection functions for detecting the utterance of voice and/or a keyword ("voice activity detection" and "keyword spotting"), recognition functions for recognizing the words spoken by the user, translating the recognized words to text and processing the intent of those words ("speech recognition" and "natural language understanding"), and a fulfillment function for executing an action requested by a command that follows the keyword. Ex. 6 (Anderson Decl.) at ¶ 37.



Pre-processing (Audio Front-End): A typical first stage in the pipeline is an "**audio front-end**," which pre-processes an audio signal received from the device's microphones to clean the signal for subsequent stages by suppressing or removing background noise and/or audio feedback from the device's speaker into the microphone, for example. Ex. 6 (Anderson Decl.) at ¶ 38.

Detection (Voice Activity Detection / Keyword Spotting): The next stages of the pipeline typically involve detection of human speech (i.e., voice) and/or specific keywords in the signal. Detection may be used as a gating mechanism for subsequent stages in the voice processing pipeline, such that subsequent processing only occurs when a keyword is detected. Ex. 6 (Anderson Decl.) at ¶ 39. As taught by the '128 and '311 patents, information obtained during the detection stage may also be used for "multi-device coordination," so that when multiple devices

(e.g., in a room together) detect a voice input, the devices can automatically negotiate leadership to determine which device should respond to the detected input. '128 patent, 1:43-2:2.

"**Voice activity detection**" (VAD) is a process by which the signal from the audio front-end is checked for human speech. As a general matter, voice activity detectors continuously monitor for and analyze an incoming audio signal as it is received, and output a prediction regarding whether speech is present. Voice activity detectors operate specifically to detect speech presence; it is a binary decision. Ex. 6 (Anderson Decl.) at ¶ 40. For example, if a user says "Hey Computer, play music," the voice activity detector would detect the presence of speech in the received audio signal.

"**Keyword spotting**," as its name suggests, operates to detect particular keywords in an audio signal. Ex. 7 (Fernández et al., *An Application of Recurrent Neural Networks to Discriminative Keyword Spotting*, ICANN 2007) ("The goal of keyword spotting is to detect the presence of specific spoken words in (typically) unconstrained speech."). A keyword may also be called a "hotword" or "wakeword," examples of which include "Hey Google" or "Hey Sonos." *See, e.g.*, '128 patent, 11:44-67 (describing uses of hotwords and wakewords for voice assistant devices); *see also* Ex. 8 (U.S. Patent No. 9,318,107, incorporated by reference in the '128 and '311 patents), 1:47-2:3, 5:2-28 (describing uses of hotwords and methods for detecting hotwords). Keywords can be used to signify a user's desire to activate voice assistant functionality. Modern approaches to keyword spotting rely on the use of neural networks, which are machine learning configurations patterned after the neurons of the human brain. Before use, the neural network is "trained" on the specific keywords of interest, after which it can predict whether an incoming audio signal contains a keyword based on characteristics of the signal. Ex. 6 (Anderson Decl.) at ¶¶ 41-42.

Keyword spotters operate specifically to detect whether one of a limited set of keywords is present in the signal based on characteristics of the audio signal; again, this is a binary decision. Ex. 6 (Anderson Decl.) at ¶ 43. For example, if a user says "Hey Computer, play music," the keyword spotter might detect that the keyword ("Hey Computer") is present, but it would not detect other words that are not keywords.

Recognition (Speech Recognition and Natural Language Understanding): After voice and/or a keyword is detected, the next step is to recognize the full set of spoken words ("speech recognition") and what they mean within the context of the system ("natural language understanding"). In other words, the recognition function is directed to the more sophisticated process of figuring out what words have been spoken and translating those words into an intent associated with the execution of a function by a device. Ex. 6 (Anderson Decl.) at ¶ 44.

"**Speech recognition**," also called "automatic speech recognition" or "ASR," involves the recognition of words in the signal and conversion of these words to text (which can then be used to process commands). Ex. 6 (Anderson Decl.) at ¶ 45. Ex. 9 (Morgan and Bourlard, *Continuous Speech Recognition*, IEEE Signal Processing Magazine (May 1995) ("The basic task of *Automatic Speech Recognition* (ASR) is to derive a sequence of words from a stream of acoustic information.")). Speech recognition employs an acoustic model to evaluate a received audio signal against thousands of known sequences of "phonemes," or units of sound that distinguish one word from another, to determine the best candidate textual words for each spoken word. Ex. 6 (Anderson Decl.) at ¶ 45. The speech recognition system then uses a language model to find the most likely word sequence from these candidate words. The language model evaluates the candidate words based on the likelihood that they will appear together in a given context—for example, it considers the likelihood that the candidate words "plan to wait" will appear together,

as opposed to the candidate words "plant weight." Ex. 6 (Anderson Decl.) at ¶ 45. Following the example described above, if a user says "play music" after the keyword, the speech recognizer would transcribe that utterance into text for further processing by subsequent stages in the voice assistant pipeline.

As compared with the detection technologies describe above (voice activity detection and keyword spotting), speech recognition is computationally expensive, requiring more memory and processing resources. Ex. 6 (Anderson Decl.) at ¶ 46. For this reason, a voice assistant pipeline may be designed such that speech recognition does not begin until a keyword has been detected.

"**Natural language understanding**" involves analysis of the recognized text corresponding to the user's utterance to determine what the user wants the voice assistant to do, typically referred to as an "intent" (e.g., a request to play music). Ex. 6 (Anderson Decl.) at ¶ 47.

Fulfillment: The final stage in the voice assistant pipeline is fulfillment. At this stage, an action is taken based on the determined user intent, and potentially based on additional information collected (for example, by the voice assistant asking the user for clarification). This stage also may involve a determination by the voice assistant as to which device(s) should fulfill the user's request. For example, as taught by the '128 and '311 patents, a user might request playback of content on a different device from the device that handles the request initially. '128 patent, 9:38-43. The user interacts with and receives responses—e.g., requests for clarification from the user or confirmation that the requested action (music playback) will be carried out—from the first device, and the user's request to play music is then fulfilled by the other device playing the requested content. '128 patent, 10:2-20, 24:61-25:7. Ex. 6 (Anderson Decl.) at ¶¶ 48-49.

B. Device Commissioning Technologies Relating to the '615 Patent

A communication network, like a Wireless Local Area Network ("WLAN"), links devices together and can connect devices to the Internet. Before a new device can join an existing

communication network it must be "commissioned" (i.e., added) to that network. The commissioning process for many networks requires any new device to have certain information before that device can join the network, including (1) the appropriate network identifier (for example, a service set identifier or "SSID") for the network, and (2) the password for that network. A variety of mechanisms can be used to supply configuration information to a new device and commission that device to a network. The novel approach in the '615 patent uses a smart wireless device that is already on the communication network to commission a new device to that network. The smart wireless device, such as a smartphone or tablet, establishes a separate short-range connection with the new device and uses that short-range path to provide configuration information to the new device. This permits the new device to join the communication network without disrupting the smart wireless device's existing connection to that same network.

III. LEVEL OF ORDINARY SKILL IN THE ART

The levels of ordinary skill at the time of the respective inventions of the Asserted Patents are set forth below:[1]

'748, '128, and '311 Patents: The parties agree that the appropriate level of skill for the '748, '128, and '311 patents is "[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." Ex. 10 (Sonos Invalidity Contentions) at 2, 4.

[1] The relevant dates of invention of the asserted patent claims are as follows: April 11, 2012 ('615 patent, all claims); July 29, 2014 ('748 patent, all claims); October 9, 2014 ('128 patent, claims 1-3, 6-8, and 11-13, and '311 patent, claims 1-3, 8-10, 12, 14-18, and 20): April 1, 2015 ('128 patent, claims 4, 9, and 14, and '311 patent, claim 19); and October 3, 2016 ('128 patent, claims 5, 10, and 15, and '311 patent, claim 11).

'615 Patent: The parties agree that the appropriate level of skill for the '615 patent is "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*., 69.

IV. LEGAL STANDARDS

A. Claim Construction

The words of a patent claim are generally given their "ordinary and customary meaning," which is "the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention" when read and understood in the context of the claims, the entire specification, and the prosecution history. *Phillips v. AWH Corp.*, 415 F.3d 1303, 1313 (Fed. Cir. 2005). "There are only two exceptions to this general rule: 1) when a patentee sets out a definition and acts as his own lexicographer, or 2) when the patentee disavows the full scope of a claim term either in the specification or during prosecution." *Thorner v. Sony Computer Entm't Am. LLC*, 669 F.3d 1362, 1365 (Fed. Cir. 2012). When a claim term is "clear on its face," the claim construction process is limited to "determining if a deviation from the clear language of the claims is specified" under the rule articulated in *Thorner*. *K-2 Corp. v. Salomon S.A.*, 191 F.3d 1356, 1363 (Fed. Cir. 1999); *Interactive Gift Express, Inc. v. Compuserve, Inc.*, 256 F.3d 1323, 1331-32 (Fed. Cir. 2001); *see also U.S. Surgical Corp. v. Ethicon, Inc.,* 103 F.3d 1554, 1568 (Fed. Cir. 1997) ("Claim construction . . . is not an obligatory exercise in redundancy.").

Extrinsic evidence may also "assist the court in determining the meaning of particular terminology to those of skill in the art of the invention." *Phillips*, 415 F.3d at 1318. But extrinsic evidence should not "contradict the meaning otherwise apparent from the intrinsic record." *Helmsderfer v. Bobrick Washroom Equip., Inc.*, 527 F.3d 1379, 1382 (Fed. Cir. 2008).

B. Indefiniteness

Invalidity of a patent claim must be proven by clear and convincing evidence. *Microsoft Corp. v. I4I Limited Partnership*, 564 U.S. 91, 95 (2011). A patent claim is not indefinite if the claim, read in light of the specification and the prosecution history, "inform[s], with reasonable certainty, those skilled in the art about the scope of the invention." *Nautilus, Inc. v. Biosig Instruments, Inc.*, 572 U.S. 898, 901 (2014). *Nautilus* "mandates clarity, while recognizing that absolute precision is unattainable." *Id.* at 910. It also recognizes that "[s]ome modicum of uncertainty . . . is the 'price of ensuring the appropriate incentives for innovation.'" *Id.* at 909 (quoting *Festo Corp. v. Shoketsu Kinzoku Kogyokabushiki Co.*, 535 U.S. 722, 732 (2002)).

C. Correcting Obvious Errors

The Commission applies the same standard as district courts to determine whether an obvious claim-drafting error should be corrected during claim construction. Specifically, obvious errors are correctable where "(1) 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." *Certain Multiple Mode Outdoor Grills and Parts Thereof*, Inv. No. 337-TA-895, Comm'n Op., 2015 WL 13817118 at *16 (Feb. 3, 2015) (citation omitted) (emphasis in original). As the Commission has explained, "absent evidence of culpability or intent to deceive, a patent claim should not be invalidated based on an obvious clerical error." *Id.* (citing *Hoffer v. Microsoft Corp.,* 405 F.3d 1326, 1331 (Fed. Cir. 2005)).

V. '748 PATENT

A. Overview of the '748 Patent

The '748 patent claims systems and methods for designating a device from a plurality of devices to perform speech recognition. *See generally, e.g.*, '748 patent, 2:36-39. This designation prevents parallel-processing issues that may otherwise arise from multiple devices attempting to

simultaneously "create, access, modify, or delete the same data." *Id.*, 3:7-13, 5:37-56. This designation is made in part based on information received from the various devices, such as a configured hierarchy based on type of device (7:37-64), audio input quality of each device's environment (7:65-8:13), the number of predetermined audio commands each device is capable of handling (8:14-25), the "context" of each device such as location, the time, applications installed on or executing on the device, or other contextual information (8:26-52); or other information related to the user's identity (8:53-62). Upon determining it is the device best suited to respond to an audio input, the designated device may send "an indication of the designation" to the other devices, and the other devices may "cease processing of audio input." The designated device is then able to perform any necessary follow-up and fulfill the user's request (or instruct another device to do so, if appropriate). *Id.*, 18:17-20, 25:43-57.

Once a device has been designated through this process, it may "utilize speech recognition techniques to determine text from obtained audio input, and compare the text to a set of predetermined audio commands … to determine whether the obtained audio input includes a predetermined audio command." *Id.*, 4:57-63. Predetermined audio commands are spoken inputs that are "associate[d] … with an action to be performed by the respective computing device," such as "make a call," "what's the time," "give me directions," or "send email." *Id.*,4:8-21, 9:61-62, 20:16-17, 24:28-30. If a predetermined audio command is detected, the device may then request any necessary additional information and/or execute the action associated with the command. *Id.*, 24:31-51.

B. Agreed Constructions

Claim Term	Agreed Construction
"the sec-end first computing device" (Claim 3)	"the first computing device"

C. "[perform/performing] speech recognition" (All Claims)

Complainant's Proposed Construction	Respondent's Proposed 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"

The term "speech recognition" has a well-understood meaning in the art, it refers to a process for transforming speech to corresponding text, as reflected in Google's proposal. Sonos's proposal, on the other hand is overly broad and would effectively write the words "speech recognition" out of the claims.

The term "speech recognition" is well-understood in the art to refer to a process for transforming speech to corresponding text. Ex. 6 (Anderson Decl.) at ¶ 52. Nothing in the specification or prosecution history indicates that the term should have anything other than its plain and ordinary meaning. Sonos's proposal is unclear and effectively excludes the plain and ordinary meaning of the term. It would improperly sweep detection techniques (voice activity detection and keyword (hotword) detection) within the term's scope (*see generally supra* Section II.A), and is redundant of other elements recited in the claims.[2] Indeed, Sonos's construction is even broader than that, seeking to ensnare ***any*** processing of "spoken audio to determine whether it includes a predetermined spoken audio command." Ex. 6 (Anderson Decl.) at ¶ 54, 58. Thus, Sonos's construction, which is overly broad in certain ways and unduly narrow in others, should be rejected

The '748 patent uses the term "speech recognition" consistent with its ordinary meaning of deriving text from speech. The '748 patent describes the claimed computing devices' "audio input modules may be operable to utilize ***speech recognition techniques to determine text from***

[2] For example, "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." '748 patent, cl. 1.

obtained audio input." '748 patent, 4:57-59 (emphasis added). A POSITA would understand from this disclosure that the term "speech recognition," as used in the '748 patent has its plain and ordinary meaning. Ex. 6 (Anderson Decl.) at ¶ 53; *supra* Section II.A. As reflected in literature disclosing the state of the art at the time of the '748 patent, Google's proposal (and the description in the '748 patent) is the well-accepted meaning of "speech recognition." *See, e.g.*, Ex. 6 (Anderson Decl.) at ¶ 52; Ex. 9 (Nelson Morgan & Hervé Bourlard, *Continuous Speech Recognition*, IEEE Signal Processing Magazine (May 1995)) at 3-4 ("The basic task of Automatic Speech Recognition (ASR) is to derive a sequence of words from a stream of acoustic information"); Ex. 11 (Keshet, Automatic Speech and Speaker Recognition (2009)) at 3 (the goal of "automatic speech recognition" is "to transcribe a recorded speech utterance into its corresponding sequence of words.").

Sonos's proposed construction is demonstrably incorrect. First, Sonos seeks to broaden "speech recognition" to include any kind of "processing" (without explaining what this term means), including keyword or hotword detection. However, a POSITA would recognize that speech recognition (transforming speech to text) differs from keyword or hotword detection, which have different uses and implications. Ex. 6 (Anderson Decl.) at ¶¶ 54-55; *see, e.g.*, Ex. 11 (Keshet, Automatic Speech and Speaker Recognition (2009)) at 3 (the goal of "automatic speech recognition" is "to transcribe a recorded speech utterance into its corresponding sequence of words."), 13 ("The core objective of a keyword spotting system is to discriminate between utterances in which a given keyword is uttered to utterances in which the keyword is not uttered."); *see also id.* at 12 (describing limitations of keyword spotting: "Most of these limitations are ***not specific to the keyword spotting problem***, and are common to ***other tasks such as speech recognition***, as pointed out in Chapter 1."); *supra* Section II.A. The distinction between "speech

recognition" and keyword or hotword detection is reflected in the specification. As noted above, the specification is unequivocal in its description of speech recognition techniques "***to determine text from obtained audio input***." '748 patent, 4:57-59. The ***only*** time the '748 patent uses the term "hotword" is in the context of monitoring for a ***predetermined audio command***, not speech recognition. *See* '748 patent at 2:66-67.

Second, and similarly, Sonos's construction conflates two separate processes performed by the claimed computing devices, only one of which is "performing speech recognition." The '748 patent explains that the computing devices may "process the audio input by performing one or more speech-to-text transformations corresponding to the received audio input" (i.e., speech recognition) or, "[i]n other examples," by performing "audio analysis to recognize and identify speech, such as performing frequency analysis to identify tonal characteristics or other sonic identifiers present in the audio input." '748 patent, 15:28-38. A POSITA reading the '748 patent would thus understand the latter to refer to processes separate from speech recognition within the context of the patent. Specifically, a POSITA would understand "performing frequency analysis to identify tonal characteristics or other sonic identifiers present in the audio input" to refer to voice activity detection (i.e., whether the audio received by the device was a voice) or keyword detection, but not speech recognition. Ex. 6 (Anderson Decl.) at ¶¶ 56-57.

Sonos's inclusion of "determin[ing] whether it includes a predetermined spoken audio command" thus cannot be correct.[3] For the foregoing reasons, "[perform/performing] speech

[3] Sonos's proposed construction is also at odds with, and undermined by, its own invalidity contentions in this Investigation. Sonos contends that "performing speech recognition" refers to "processing spoken audio," but simultaneously argues in its contentions that "speech recognition" is not sufficiently described in the '748 patent. Ex. 10 (Sonos Invalidity Contentions) at 3 ("[T]he specification lacks enablement and written description support for this recited sequence. Indeed, the specification only mentions the term 'speech recognition' twice in a more general sense. *See, e.g.*, *id.*, 4:57-63, 10:21-25. There is no disclosure of a first computing device 'determin[ing] whether [it] should perform speech recognition on the spoken

recognition" should be given its ordinary meaning, and Sonos's overly broad proposal should be rejected.

D. "a command to temporarily refrain from performing speech recognition" (All Claims)

Complainant's Proposed Construction	Respondent's Proposed Construction
No construction necessary (plain and ordinary meaning)	"a message instructing the second computing device to temporarily cease performing any and all speech recognition"

This term should be accorded its plain and ordinary meaning, consistent with the intrinsic record. Sonos's proposed construction substitutes some words of the claim and adds others (i.e., "any and all"). Sonos's changes do not, on their face, appear to alter the ordinary meaning of the disputed claim phrase—e.g., substituting "message instructing" for "command," and "cease" for "refrain." But the original claim language itself is clear, and Sonos's substitutions are therefore unnecessary and disfavored, particularly if Sonos advocates an interpretation of its construction that differs from the ordinary meaning articulated by Google below. Sonos's other addition—"any and all"—merely adds words that lack support in the intrinsic evidence (through lexicography or disclaimer) and should also be rejected.

The term "a command to temporarily refrain from performing speech recognition" appears in each of the independent claims of the '748 patent. For example, claim 1 recites "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***." '748 patent, cl. 1 (emphasis added). Claims 7 and 11 include similar limitations. The claim language is clear on its face. It uses common English words and employs them using their standard

audio input,' much less performing additional functions (i.e., send, perform, and determine) based on such a determination.").

meanings. Each independent claim requires, "responsive to determining that the first computing device should perform speech recognition on the spoken audio input," "send[ing], [by the first computing device and] to at least a second computing device from a plurality of computing devices, a command to temporarily refrain from performing speech recognition." '748 patent, cls. 1, 7, 11. The plain meaning of this claim language is directed to the operation of the claimed first computing device, which sends a message to at least the second computing device that at least the second computing device should not perform speech recognition for a limited amount of time.

The specification confirms the claim language's ordinary meaning to a POSITA. Ex. 6 (Anderson Decl.) at ¶¶ 61-62. The '748 patent explains that, "[i]n general, techniques of the present disclosure may enable one or more computing devices from among a plurality of computing devices to designate a particular computing device to monitor for and/or act on audio input . . . ." '748 patent, 2:36-41. Once that device is designated, the other devices may "temporarily cease to monitor for audio input." *Id.*, 2:52-57. One way this can be accomplished is for the designated computing device to send instructions (a command) to the other, non-designated computing devices informing them that they should temporarily cease monitoring for audio input. *Id.*, 9:36-41 ("In some examples, the designated particular computing device may output instructions to computing devices other than the particular computing device to cause the respective computing devices other than the particular computing device to temporarily deactivate respective audio input modules.").

A POSITA would have no difficulty understanding the meaning of the disputed claim language in view of these specification disclosures. Ex. 6 (Anderson Decl.) at ¶¶ 62-63.

Sonos's construction changes some words in the claim and adds others without support in the claims or intrinsic record, without altering the meaning of the disputed phrase in any apparent

way. Specifically, Sonos changes "command" to "message instructing," "refrain" to "cease," and "speech recognition" to "any and all speech recognition." It is not clear how, if at all, Sonos intends those changes to affect the claim scope, or how Sonos's changes could provide a POSITA with any added clarity as to the meaning of the claim language. Claim construction is "not an obligatory exercise in redundancy"; there is no reason to construe claim language that is clear on its face, particularly in a way that does not add any clarity or affect claim scope. *U.S. Surgical Corp. v. Ethicon, Inc.*, 103 F.3d 1554, 1568 (Fed. Cir. 1997).

For these reasons, Google's proposal of plain and ordinary meaning should be adopted.

E. "predetermined spoken audio command" (All Claims)

Complainant's Proposed Construction	Respondent's Proposed Construction
No construction necessary (plain and ordinary meaning)	"any audio input usable to initiate interaction with a computing device, such as a spoken hotword"

"Predetermined spoken audio command" requires no construction. It has a plain and ordinary meaning to a POSITA in light of the specification. A predetermined spoken audio command is a spoken utterance that includes a directive for the claimed computing device to execute an associated action. Ex. 6 (Anderson Decl.) at ¶ 66. Sonos's construction for this phrase is directed to "initiat[ing] interaction" with a computing device, and would therefore vitiate the requirement that the predetermined spoken audio command actually include a ***command***, i.e., an instruction to execute an action, as recited in the asserted claims. Only Google's proposal is consistent with the intrinsic evidence.

The claims require the "predetermined spoken audio command" to include a directive that can be executed by the device. Ex. 6 (Anderson Decl.) at ¶¶ 67-68. Each of independent claims 1, 7, and 10 recites "determin[e/ing], 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***." '748 patent, cls. 1, 7, 10.

The specification consistently and repeatedly associates predetermined audio commands with corresponding actions. *See, e.g.*, '748 patent, 1:15-18 ("In some cases, a computing device may monitor an environment for audio input that includes a predetermined audio command and, if the predetermined audio command is detected, perform a corresponding action."), 4:8-10 ("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").

Indeed, the specification is replete with specific examples of predetermined spoken audio commands and their associated actions. The '748 patent explains that "[o]ne or more of audio input modules 6 may associate a predetermined audio command with an action to be performed by the respective computing device." '748 patent, 4:8-10. For example, predetermined audio commands can include "answer" (to answer a telephone call), "navigation" or "directions" (to request directions), "send email," "send text message," and other directive phrases. *Id.*, 4:14-21. Rather than simply "directions," the phrase "[g]ive me directions" may also be a predetermined spoken audio command. *Id.*, 9:61-62 ("The user may speak a predetermined audio command (e.g., 'Give me directions')."). In another example, "the audio input includes the predetermined audio command 'what's the time." *Id.*, 20:15-18. "Go silent" is another example. *Id.*, 24:36-39. The specification also provides an example where "the predetermined audio command may be the word 'play' and the subsequent audio input may be the title of a song, TV show, or other media content item." *Id.*, 24:41-46.

Sonos seeks to read the word "command" out of the claims entirely. By defining "predetermined spoken audio command" to cover "***any*** audio input usable to ***initiate interaction*** with a computing device, such as a spoken hotword," Sonos effectively replaces a "command" in the claims with a keyword (e.g., "spoken hotword"). This cannot be correct, at least because many "commands" expand beyond the bounds of a hotword. As discussed above, the specification discloses predetermined audio commands associated with a variety of actions to be carried out by the device beyond just entry into an audio input mode (as would be the case with a hotword), including answering telephone calls and sending messages. *See, e.g.*, '748 patent, 4:8-30. Sonos's proposed construction would exclude these disclosed "actions," thereby unduly restricting and otherwise changing the meaning of the claim term, and should be rejected.

Sonos's construction is unduly narrow and would exclude clear "commands," as described above. Excluding such embodiments is disfavored and, for the reasons described above, the plain and ordinary meaning of "predetermined spoken audio command" should govern.

F. "[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" (All Claims)

Complainant's Proposed Construction	Respondent's Proposed 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"

This claim term, as written, is clear. The term appears in each of the independent claims of the '748 patent. For example, claim 1 recites: "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." '748 patent, cl. 1. Claims 7 and 11 include similar language. '748 patent, cls. 7, 11.

The parties do not dispute this portion of the claim phrase: "determining . . . whether the first computing device should perform speech recognition on the spoken audio input." The dispute centers on Sonos's wholesale replacement of the claim language "based at least in part on information received from the one or more additional computing devices" with its proposed construction of "without any of the plurality of computing devices processing the spoken audio input to determine whether it includes a predetermined audio command" (i.e., "speech recognition," as it has construed that term). Sonos's construction would alter the plain meaning of the claim in two material ways: (1) it would broaden the claim in one respect by eliminating the requirement that the determination of "whether the first computing device should perform speech recognition" be "based at least in part on information received from the one or more additional computing devices"; and (2) it would narrow the claim in another respect by adding a requirement that this determination be done without any other device of the plurality of computing devices performing additional "processing" (i.e., "speech recognition").

Because the term is clear on its face, Sonos must demonstrate that the patentee either intended this limitation to have a special meaning or disavowed claim scope. *K-2 Corp.*, 191 F.3d at 1363. Neither occurred here. The specification explains that "one or more computing devices from a plurality of computing devices may designate a particular computing device from among the plurality of computing devices to process audio input to determine whether the audio input includes a predetermined audio command." '748 patent, 6:61-66. The computing devices may determine which of them should be designated as the "particular computing device" using an arbitration process, or each can make the decision individually. *Id.*, 7:30-36.

The determination can be "based at least in part on" a variety of criteria, such as a configured hierarchy based on type of device (7:37-64), audio input quality of each device's environment (7:65-8:13), the number of predetermined audio commands each device is capable of handling (8:14-25), the "context" of each device such as location, the time, applications installed on or executing on the device, or other contextual information (8:26-52); or other information (8:53-62). The claim language ("based at least in part on information received from the one or more additional computing devices") is consistent with these examples, and the plain meaning of the claim language would be readily understood by a POSITA.[4] Ex. 6 (Anderson Decl.) at ¶ 72.

Sonos's construction impermissibly broadens the scope of the claims by removing the requirement that the "determining" step must be "based at least in part on information received from the one or more additional computing devices." It is well-established that "each element contained in a patent claim is deemed material to defining the scope of the patented invention." *Bicon, Inc. v. Straumann Co.*, 441 F.3d 945, 950 (Fed. Cir. 2006); *see also Merck & Co. v. Teva Pharms. USA, Inc.*, 395 F.3d 1364, 1372 (Fed. Cir. 2005) ("a claim construction that gives meaning to all the terms of the claim is preferred over one that does not do so").

Sonos's construction also narrows the scope of the claims by adding a negative limitation: that the "determining" step be performed "without any of the plurality of computing devices processing the spoken audio input to determine whether it includes a predetermined audio command." This aspect of Sonos's construction—which would exclude from the claim scope a system in which the component devices process at least some spoken audio input even if they are

[4] Whereas the above examples describe designating a device "to process audio input," the claims recite determining "whether the first computing device should perform speech recognition." As described above in Section V.C, speech recognition is a form of audio processing.

not the designated device[5]—finds no support in the intrinsic record. Moreover, negative limitations must find support either in "the words of the claim" or through an "express disclaimer or independent lexicography in the written description that would justify adding that negative limitation." *Omega Eng'g, Inc, v. Raytek Corp.*, 334 F.3d 1314, 1323 (Fed. Cir. 2003). There is no such support here.

Sonos's construction would require the claims' determination "whether the first computing device should perform speech recognition on the spoken audio input" to be performed without any other device performing "speech recognition" (as defined by Sonos). However, nothing in the intrinsic evidence precludes the non-designated computing devices from beginning speech recognition before a first computing device determines that it should do so; the patentee left those implementation details open. There is nothing precluding such operation in the specification, which explicitly permits all computing devices to monitor the environment and perform some processing on the received audio input. For example, where the designation decision is based on "a level of audio input quality associated with one or more of computing devices 2," the '748 patent explains that "one or more of computing devices 2 may obtain audio input from the environment . . . and communicate with others of computing devices 2 to compare the quality of the audio input and/or of speech included therein." *Id.*, 8:2-7. That communication would require some processing, but there is no reason that processing could not include determining whether the audio input includes a predetermined audio command. Ex. 6 (Anderson Decl.) at ¶ 74. Nor does the specification foreclose speech recognition from initiating at other devices before they receive an instruction from the first computing device to refrain from doing so after the first computing

[5] This restriction is also incorrect for the reasons addressed above with respect to the term "[perform/performing] speech recognition." *See supra* Section V.C.

device determines that it will perform speech recognition and, correspondingly, execute an action associated with the predetermined spoken audio command. The claims (consistent with the specification) are focused on the operation of the first electronic device in deciding to perform speech recognitions, and the steps the first electronic device takes (including instructions issued to other devices) in response to making that determination.

Further, even when the non-designated devices are instructed not to perform speech recognition, the focus of the claim is on the issuance of the instruction and not on how exactly and when any device that receives the instruction operates in response. As discussed above with respect to the term "a command to temporarily refrain" in Section V.D, in some embodiments the designated device may send a command to the other devices to refrain from performing speech recognition on the spoken audio input. When the designated device sends an instruction to the non-designated devices to "temporarily deactivate respective audio input modules," the '748 patent does not specify that the deactivation must be immediate or that it must occur before any of the spoken audio input is processed. *Id.*, 9:36-50. Therefore, there is no basis to construe the claims to require that the determination whether to perform speech recognition must be made "without any of the plurality of computing devices processing the spoken audio input to determine whether it includes a predetermined audio command."

For the foregoing reasons, this term should also be accorded its plain and ordinary meaning, and Sonos's construction should be rejected.

VI. '128 AND '311 PATENTS

A. Overview of the '128 and '311 Patents

The '128 and '311 patents are related and share a common specification. *See* '311 patent, p. 2 (Related U.S. Application Data). The patents describe devices with "voice assistants" that allow users to "implement voice-activated functions" such as "initiate a phone call, conduct a

restaurant search, start routing on a map, create calendar events, add a post to a social network, recognize a song, and complete many other tasks." '128 patent, 1:34-37. As the patents explain, a location ("e.g., a room or space within a home") may include multiple such devices, and it is "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." *Id*., 1:43-53.

To achieve these goals, the patents describe methods and systems for "multi-device coordination" to "negotiate leadership amongst multiple devices for response to a speaker." *Id.*, 1:57-2:2. This coordination can be performed by comparing values associated with detection of the voice inputs at the respective devices, which may be affected by multiple factors, including "surrounding noise, capabilities of the device, distance and position of the device relative to the user, position of the device within the space, and so on." *Id*., Figs. 4A-4B and 6, 20:45-67, 28:60-29:8. The patents describe that each device determines "a level of quality" of the detected voice input, which "may be expressed as a score or some other value," and communicates that value or score to other devices. *Id*., 21:1-4, 21:14-22, 29:9-31. The values or scores of the respective devices are then compared, and based on that comparison, a determination regarding whether a device should respond to the user's voice input or forgo responding is made. *Id*., 21:58-22:1, 29:32-30:12.

The above-described comparison based on values or scores determined based on detection of a voice input is reflected, for example, in claim 1 of the '128 patent, which recites "detecting a voice input" and "determining a quality score for the detected voice input." '128 patent, cl. 1. The patents also describe how devices might use additional information gained from speech recognition to make a separate determination whether to respond based on the content of the user's request. For example, claim 5 of the '128 patent recites "recognizing a command in the voice

input" and responding if "the type of the command is related to the first electronic device." '128 patent, cl. 5.

A device that responds to a detected voice input outputs audible or visual responses to the user, for example to request additional information or to confirm that the user's request has been fulfilled. *Id*., 10:2-20, 24:61-25:7. However, a user's request might be fulfilled by a different device from the device that interacts with the user; in other words, a device could fulfill the user's request even if it forgoes responding to the detected voice input based on the above-described comparison of values or scores. *Id*., 9:38-43, 24:45-51.

B. "other devices of the plurality of electronic devices" / "quality scores generated by the other devices of the plurality of electronic devices" ('128 All Claims)

Complainant's Proposed Construction	Respondent's Proposed 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"

The independent claims of the '128 patent recite "a first electronic device of a plurality of electronic devices." Sonos has not identified the "plurality" term for construction, and there is no dispute that the ordinary meaning of "plurality" is "at least two." Instead, the parties' dispute centers on the "other devices"/"quality scores" terms, which appear in the following limitations:

> 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

'128 patent, claims 1, 6, 11 (emphases added).

Specifically, Sonos contends that the terms "other devices" and "quality scores" require two or more devices and two or more scores (thereby excluding a system having only two devices—a first electronic device and an "other device"). This is incorrect. The meaning of these terms is clear in the context of the claims and the specification. "Other devices" may refer to one or more "other devices;" it does not require two or more devices. This principle also applies to "quality scores," which may refer to a single "quality score" (for example, when there is only one "other device") or multiple "quality scores" (when there are multiple "other devices").

The ordinary meaning of the term "plurality," recited in the asserted claims of the '128 patent, "simply means two or more." *Cheese Sys., Inc. v. Tetra Pak Cheese & Powder Sys., Inc.*, 725 F.3d 1341, 1348 (Fed. Cir. 2013); *see also York Prods., Inc. v. Cent. Tractor Farm & Family Ctr.*, 99 F.3d 1568, 1575 (Fed. Cir. 1996). By using this term, the asserted claims are directed to a system of "two or more" electronic devices. 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." Ex. 6 (Anderson Decl.) at ¶ 79. The Federal Circuit has repeatedly cautioned that plural terms (and whether those terms can cover a single device) should be construed in light of the context in which they appear. 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"). *See Versa Corp. v. Ag-Bag Int'l Ltd.*, 392 F.3d 1325, 1330 (Fed. Cir. 2004) (holding "means for creating air channels" did not require a supporting structure that made

multiple channels); *Omeprazole Patent Litigation*, 84 F. App'x 76, 80 (Fed. Cir. 2003) (non-precedential) (claims reciting "one or more layers of materials" did not require two or more materials because "materials" reflects an attempt to achieve grammatical consistency with the tense of "layers").

Indeed, the Federal Circuit's decision in *Dayco Prods., Inc. v. Total Containment, Inc.*, which addressed claims reciting a "plurality" and plural claim elements within that plurality, is on all fours with the facts here. 258 F.3d 1317, 1328 (Fed. Cir. 2001). In *Dayco*, the Federal Circuit found that the district court erred in construing "a ***plurality*** of outwardly directed projections with ***recesses*** there between" to require at least three projections because the recitation of "recess***es***" required at least two recesses, and thus implicitly three projections. *Id.* at 1325, 1327-28 (emphases added). As the court explained, the term "plurality" "refers to two or more items, absent some indication to the contrary," and the term "recesses" in the context of the claim meant either a single recess or a plurality of recesses. *Id.* at 1327-28;[6] *see also Certain Elec. Devices, Including Handheld Wireless Commc'ns Devices*, Initial Determination, Inv. No. 337-TA-667 (Sept. 21, 2009) (in claims reciting a "plurality of processor devices" including an "owner processor device" and "other processor devices," finding that "other processor devices" may refer to a single processor device when the "plurality of processor devices" includes only two processors).

Like the claim at issue in *Dayco*, the '128 claims recite a plurality of electronic devices including a "first electronic device" and "other device***s***" (with corresponding "quality score***s***"), and the "other devices" and "quality scores" can refer to a single other device. If the '128 patent inventors had wanted to ensure a minimum of three devices (and quality scores) in a system—i.e.,

[6] The Federal Circuit noted that, had the patentees wanted to claim two or more recesses, they could have recited "a plurality of recesses" in a similar manner to its recitation of "a plurality of … projections." *Id.* at 1328.

a first electronic device and at least two "other devices"—they could have claimed a "plurality" of other devices and a "plurality" of quality scores. They did not do so.

The '128 specification confirms that the claimed "other devices" or "quality scores" can refer to a single device or score in a two-device system:

> 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, Fig. 6 (tracking elements of claim 1), 28:63-29:5 (emphases added). Here, "an electronic device" would refer to the claimed "first electronic device," and therefore "two or more electronic devices" would contemplate a single "other device." Google's proposed construction ensures that the Figure 6 embodiment is properly included within the claims' scope, as there is no evidence that the inventors desired to exclude this embodiment. *See GE Lighting Sols., LLC v. AgiLight, Inc.*, 750 F.3d 1304, 1311 (Fed. Cir. 2014) ("where claims can reasonably [be] interpreted to include a specific embodiment, it is incorrect to construe the claims to exclude that embodiment, absent probative evidence on the contrary"); *Oatey Co. v. IPS Corp.*, No. 07-1214 (Fed. Cir. Jan. 30, 2008) (similar).

For these reasons, Google's construction should be adopted, and Sonos's attempt to limit the claim scope should be rejected.

C. "second values … from other electronic devices" / "second values from other electronic devices of the plurality of electronic devices" ('311 All Claims)

Complainant's Proposed Construction	Respondent's Proposed 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"

As with the '128 claims addressed in Section VI.B, *supra*, independent claims 1 and 10 of the '311 patent recite "a first electronic device of a plurality of electronic devices." Independent claim 16 of the '311 patent recites "a plurality of electronic devices." Again, Sonos has not identified the "plurality" term for construction, nor contested its ordinary meaning of "at least two." The parties' dispute thus centers on the following limitations:

> receiving ***second values associated with the voice input from other electronic devices*** of the plurality of electronic devices (claims 1 and 10)
>
> receive respective ***second values from other electronic devices of the plurality of electronic devices*** (claim 16)

'311 patent, cls. 1, 10, 16 (emphasis added).

For the same reasons described in Section VI.B, the meaning of these terms is clear: "other electronic devices" may refer to one or more devices, and "second values" may refer to one or more "second values." Sonos's constructions would exclude two-device configurations, are unsupported by the intrinsic evidence, and contravene the relevant legal authority.

D. "forgo outputting an audible response" / "forgoing outputting a response" ('128 All Claims) / "[forgoing / forgo] responding" ('311 Claims 2, 3, 12)

Complainant's Proposed Construction	Respondent's Proposed Construction
No construction necessary (plain and ordinary meaning)	"[refrain]/[refraining] from outputting any and all [audible] responses"

The meaning of the forgo/forgoing terms is clear, and no construction is necessary. Sonos nevertheless offers a construction that improperly seeks to inject the requirement that the device forgo "any and all" responses. Sonos's added restriction highlights the parties' dispute regarding what qualifies as "responding" to a detected voice input—under Sonos's interpretation, a device that "forgo[es]" responding cannot perform any action whatsoever relating to a detected voice input, including "fulfilling" a user request. This undue restriction lacks support in the claims and is inconsistent with the specification.

The forgo/forgoing terms appear in the following limitations of each independent claim of the '128 patent:

> 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***;
>
> 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 device, ***forgoing outputting a response to the detected voice input***.

'128 patent, cls. 1, 6, 11 (emphases added). Claims 2, 3, and 12 of the '311 patent also recite "forgo[ing] responding" to the detected voice input. As the plain language of the claims makes clear, if a device in the plurality of electronic devices is determined to not have the highest quality score (or, in the context of the '311 patent, to not have a higher first value), then the device does not output an audible and/or visual response to the detected voice input. A POSITA would understand this language as written, and therefore there is no need to construe it. Ex. 6 (Anderson Decl.) at ¶ 94.

The specification confirms that the claimed "forgo[ing]" is tied to a response ***to the detected voice input***. For example, the specification describes that, "[i]f its score is not the best (e.g., highest) amongst the scores, the electronic device ***forgos outputting responses…to the voice input***." '128 patent, 21:61-65 (emphasis added); *see also id.*, 23:4-7, 29:59-30:12. Examples 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.*, 29:32-54. In contrast, a device that forgoes responding to a detected voice input may, for example, power down a speaker or associated amplifier. *Id.*, 21:61-65, 23:4-7, 29:67-30:1. In this manner, the voice assistant devices

determine a "leader" that will be responsible for responding to the user's voice inputs, and in order to reduce user confusion, other devices that are not the "leader" forgo responding to the user. *Id.*, 1:43-53.

The specification also describes the distinct process of fulfillment, by which a device takes an action or operation in response to an instruction relating to the user's request. For example, the specification describes that a user may make "a request for information or for performance of an action or operation," and "[t]he electronic device 190 ***responds to the request with voice message responses to the user***, where the response may include, for example, requests for additional information to ***fulfill the request***, confirmation that the request has been ***fulfilled***, notice that the request cannot be ***fulfilled***, and so forth." '128 patent, 10:2-20 (emphases added); *id*., 24:61-25:7 (a device may determine a request from a voice input and "generate ***a response that is responsive to the request*** (e.g., an audible acknowledgement of the request, an audible inquiry to the user for more information needed to ***fulfill the request***).") (emphases added). Thus, the specification distinguishes responses to a detected voice input (e.g., requests for additional information or confirmation messages) from acts devices might take to fulfill an action requested within that voice input.

As described in the specification, fulfillment need not be performed on a device that detects a voice input; it may take place on a device other than the device that interacts with the user. For example, "the user can issue a media play request by saying to the Wi-Fi speaker 'OK Google, play cat videos on my Living room TV.' The voice assistant service then ***fulfils the media play request*** by playing the requested media content on the requested device using a default or designated media application." '128 patent, 9:38-43 (emphasis added). In this case, if a device other than "Living room TV" detected the voice input, the "response" to the detected voice input

is the instruction sent from the detecting device to the Living room TV (i.e., the fulfilling device) instructing that TV to perform the requested action; the fulfillment performed by the Living room TV (i.e., playing cat videos) would not be the "response to the detected voice input." *Id*., 24:45-51 ("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…***response [which] may include a voice output and/or instructions to another device to perform an operation***.") (emphasis added). Thus, the '128 and '311 patents describe responding to a voice input as separate and distinct from fulfillment, and the claims do not preclude a device that forgoes outputting a response to a voice input from fulfilling an action based on an instruction from the device that does respond to the detected voice input.

By injecting the requirement that a device "refrain from outputting ***any and all*** responses," Sonos's proposed construction would broaden the scope of the claimed "response[s]" to encompass fulfillment (or any action whatsoever).[7] As described above, this restriction is contrary to the intrinsic evidence. The patents contemplate that a device may respond to a detected voice input by issuing instructions to other devices to fulfill a user's request, but does not describe the operations or actions taken by the other devices in response to those instruction as "responding to the voice input." '128 patent, 9:38-43, 24:45-51. Instead, the patents refer to that those actions as "fulfill[ment]." *Id*., 10:2-20, 24:61-25:7. Sonos's proposal would preclude devices that "forgo" responding from "fulfill[ing]" such actions, contrary to the specification's teachings.

[7] Sonos also replaces the claim term "forgoing" with "refraining from," which is unnecessary because the meaning of "forgoing" is clear based on its use throughout the specification. *See, e.g.*, '128 patent, 21:58-65, 22:66-23:7, 29:59-30:12. In contrast, neither the word "refrain" nor the phrase "refraining from" is used in the specification.

Sonos's proposal is also improper because it would preclude other device operations that are contemplated by the patents (separate from fulfillment of a user's request). The patents describe that all devices may visually indicate that they are listening to a detected voice input, while only the device with the highest quality score will respond to that input. *See* '128 patent, 23:35-45 ("In some implementations, ***each of the electronic devices 1002 outputs a visual indication*** (e.g., a predefined LED light pattern) ***that it is listening*** whenever it is listening to a user (e.g., receiving and detecting the user's voice input), and then ***just the leader device outputs the response to the voice input*** per the leadership determination.") (emphases added). Sonos's attempt to exclude such a visual indication that a device is listening to a detected voice input, but not yet responding, underscores that its proposed construction is incorrect.

No construction of the "forgo/forgoing" terms is necessary; these terms should be given their plain and ordinary meaning. Sonos's proposal should be rejected because it imports limitations into the claims and is based on an interpretation of "response to the detected voice input" that is inconsistent with the intrinsic record.

E. "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" ('128 Claims 1, 2, 3, 5)[8] / "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" ('311 Claim 2)

Complainant's Proposed Construction	Respondent's Proposed Construction
No construction necessary (plain and ordinary meaning)	Indefinite

The meaning of the "in accordance with a determination…" terms is clear from the context of the claims and the specification, and no construction is necessary. In asserting indefiniteness,

[8] Claims 2, 3, and 5 of the '128 patent depend from claim 1.

Sonos contends the claims lead to a non-sensical result, asserting that "it is impossible for an 'electronic device' to perform both of these method steps for a given voice input." *See* Ex. 10 (Sonos Invalidity Contentions) at 6-8. That is simply incorrect.

'128 patent claim 1 and '311 patent claim 2 recite methods in which two of the steps are presented as alternatives. Whether one or the other is taken depends on what happens at earlier conditional steps of the claims, i.e., the claim terms in dispute. Thus, for example, the step of "***outputting*** an audible and/or a visual response to the detected voice input" is performed "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." '128 patent, cl. 1. And "***forgoing outputting*** a response to the detected voice input" is alternatively performed "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." *Id*. Claim 2 of the '311 patent recites similar steps using a similar framework.[9] '311 patent, cl. 2. Nothing about these conditional steps renders these claims indefinite. Rather, the plain meaning of the disputed phrases in the context of the claims and specification is clear, and thus Google's proposals should be adopted.

First, method claims may include conditional steps. *Cybersettle, Inc. v. Nat'l Arbitration Forum, Inc.*, 243 F. App'x 603, 607 (Fed. Cir. 2007) ("It is of course true that method steps may be contingent."); *see also Hytera Commc'ns Co. v. Motorola Sols., Inc.*, 841 F. App'x 210 (Fed. Cir. 2021) (affirming validity over prior art of claim with limitation "written in conditional

[9] Google understands Sonos's position to be the same for the '128 and '311 terms. For purposes of this Markman Brief, Google uses the "quality score" and "highest" / "not highest" language from claim 1 of the '128 patent. Claim 2 of the '311 patent uses the same framework (i.e., "value" and "higher" / "not higher"), and the arguments presented here apply equally to that claim.

language, with two alternative conditions and corresponding responses"); *Interdigital Tech. Corp. v. Lenovo Holding Co., Inc.*, No. CV 19-1590-LPS, 2021 WL 1856937, *5 (D. Del. May 10, 2021) (construing alternative conditional limitations).

Second, the claim language and specification confirm the plain meaning of the disputed phrases. As the specification explains, "if" has the same meaning as the claim phrase "in accordance with a determination that." *Id*., 35:6-9 ("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.") (emphases added). The claim's use of "in accordance with a determination that" thus provides alternative conditions precedent, which would have been readily understood by a POSITA within the context of the patents. Ex. 6 (Anderson Decl.) at ¶¶ 111, 113-17.

The specification also explains that the conditional aspect of the disputed claim terms derives from the very purpose of the claimed invention. The '128 and '311 patents present solutions to problems that arise when a single location (e.g., a room) includes "multiple devices that include voice assistant systems and one or more users of such devices." '128 patent, 1:43-45, 59-63. To address these issues, the patents disclose methods for "multi-device coordination" to determine "a leader amongst the voice assistant devices that is responsible for responding to the user's voice inputs, in order to reduce user confusion." *Id*. at 1:50-53, 59-63. In this context, a POSITA would expect the claimed invention to encompass steps that a device takes if it is the "leader," and steps that it alternatively takes if it is not the "leader." Ex. 6 (Anderson Decl.) at ¶ 118.

The specification thus mirrors the conditional framework of the claims, consistently using the phrase "in accordance with a determination that" to describe the alternative functionalities. For

example, the specification describes that, ***"[i]n accordance with a determination that*** the quality score generated by the first electronic device ***is the highest***…the electronic device ***outputs*** (1210) an audible response and/or a visual response (e.g., a notification or a LED light pattern) to the detected voice input." '128 patent, 29:32-40 (emphases added). The specification describes the alternative using the same language: "***[i]n accordance with a determination that*** the quality score generated by the first electronic device is ***not the highest***…the electronic device 190 ***forgoes outputting*** a response to the detected voice input." *Id*., 29:59-63 (emphases added). The specification confirms the conditional nature of this step through its use of the term "if" in the next sentence: "***If*** a device determines that its score is not the highest amongst the devices with respect to the voice input (e.g., devices 1002-2 and 1002-3), the device does not output a response to the voice input." *Id*., 29:63-67 (emphasis added); *see also id.* at 21:58-22:10 (using "if" to describe the same operations). These descriptions confirm to a POSITA that the disputed limitations recite conditional steps. Ex. 6 (Anderson Decl.) at ¶ 119.

In asserting that these limitations are indefinite, Sonos ignores both their plain meaning in the context of the patents and well-settled law. First, Sonos's indefiniteness argument is premised on the notion that the claims lead to a non-sensical result by requiring a determination that the first electronic device's quality score is both "the highest" ***and*** "not the highest," and that the first electronic device must therefore both output a response ***and*** forgo outputting a response. As explained above, a POSITA would readily understand the conditional framework of the claims, particularly in light of the specification's teachings. Ex. 6 (Anderson Decl.) at ¶¶ 111-18. A POSITA would also readily appreciate that both conditions stated in the '128 and '311 claims cannot be met simultaneously, and therefore would not interpret the claims as Sonos suggests. Ex. 6 (Anderson Decl.) at ¶ 120. *AIA Eng'g Ltd. v. Magotteaux Int'l S/A*, 657 F.3d 1264, 1276 (Fed.

Cir. 2011) ("We strive, where possible, to avoid nonsensical results in construing claim language.").

Second, Sonos's position contravenes Federal Circuit authority. *See, e.g.*, *Hytera*, 841 F. App'x at 215-218 (discussing *Lincoln Nat. Life Ins. Co. v. Transamerica Life Ins. Co.*, 609 F.3d 1364 (Fed. Cir. 2010). Following that authority, the argument Sonos makes here was rejected in *Interdigital*, which involved a claim with conditional, "mutually exclusive alternative[]" steps: "[s]tep two recites a ***condition*** in which an explicit allocation ***is*** received…step three recites the ***condition*** in which an explicit allocation ***is not*** received." 2021 WL 1856937 at *5 (emphases original). The court held that "[i]t would be impossible for a system both to receive and not receive an explicit allocation at any given time" and, "to infringe[,] a system must be capable of addressing each alternative, regardless of which alternative occurs at any particular point." *Id*. (citing *Hytera* and *Lincoln Nat. Life Ins. Co. v. Transamerica Life Ins. Co.*, 609 F.3d 1364 (Fed. Cir. 2010)). The same logic applies here.

Because a POSITA would readily understand these terms, they should be given their plain and ordinary meaning, and Sonos's indefiniteness argument should be rejected.

F. "cause the electronic device" ('128 Claims 11, 12, 13, 15)[10]

Complainant's Proposed Construction	Respondent's Proposed Construction
No construction necessary (plain and ordinary meaning)	Indefinite

The phrase "cause the electronic device" has a plain meaning in light of the '128 patent claims and specification. Sonos's indefiniteness position is premised on an assumption that the preamble of claim 11 does not specify whether "the electronic device" in the disputed phrase is

[10] Although this phrase does not appear in claims 12 and 13, a construction would affect these claims because they depend from claim 11.

the same device as the "first electronic device." *See* Ex. 10 (Sonos Invalidity Contentions) at 7-8.[11] It is the same device, as the '128 patent confirms. Ex. 6 (Anderson Decl.) at ¶ 124. No construction is necessary.

First, the preamble of claim 11 recites "[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 [the operations recited in the claim]." '128 patent, cl. 11 (emphases added). The only device of "a plurality of electronic devices" that claim 11 specifies is "the first electronic device." *Id*. Accordingly, a POSITA would understand that the only possible antecedent basis for "the electronic device" recited in claim 11 is the "first electronic device." Ex. 6 (Anderson Decl.) at ¶ 125. *See Joy MM Delaware Inc. v. Cincinnati Mine Mach. Co.*, No. CIV.A. 09-1415, 2010 WL 3984822 (W.D. Pa. Oct. 9, 2010), *aff'd*, 497 F. App'x 970 (Fed. Cir. 2012) (claim's recitation of "a chain and flight assembly" and "said ***conveyor*** chain and flight assembly" did not render it indefinite where the claim only referred to one "chain and flight assembly").

This understanding is confirmed by the context of claim 11, which recites "instructions" that are executed by "a first electronic device." The claim further recites that these instructions, when executed, cause "the electronic device" to perform a number of operations that the claim confirms are carried out by the "first electronic device." '128 patent, cl. 11 ("the quality score generated by the ***first electronic device***") (emphasis added). Because "the electronic device"

[11] Sonos asserts in its contentions that claims 11, 12, 13 and 15 are indefinite "[f]or at least this reason," but has not responded to Google's request for identification of any additional reasons for its indefiniteness position. Ex. 10 (Sonos Invalidity Contentions) at 8.

performs the recited operations based on execution of instructions, a POSITA would understand "the electronic device" to be the same device that executes those instructions (i.e., the "first electronic device"). Ex. 6 (Anderson Decl.) at ¶ 126.

Second, the specification confirms that "the electronic device" of claim 11 is the recited "first electronic device." For example, Figure 6 of the '128 patent illustrates a series of operations that correspond to those recited in claim 11, and shows that those operations are performed "[a]t a first electronic device of a plurality of electronic devices." '128 patent, Fig. 6. At the same time, the specification describes Figure 6 as showing a method that is "implemented at an ***electronic device***" of two or more electronic devices. *Id.*, 28:65-30:12 (emphasis added). Based on this description, a POSITA would understand that the electronic device recited in claim 11, which is caused to perform the same operations illustrated in Figure 6, is the same as the "first electronic device" recited in claim 11. Ex. 6 (Anderson Decl.) at ¶¶ 127-28. The intrinsic record supports only this interpretation. *See Eidos Display, LLC v. AU Optronics Corp.*, 779 F.3d 1360, 1364-66 (Fed. Cir. 2015) (claim term not indefinite where the specification only described one type of structure).

For the same reasons, Sonos's indefiniteness position regarding "cause the electronic device" is inconsistent with the intrinsic record and should be rejected. No construction is necessary, and this term should be given its plain and ordinary meaning.

G. "the second scores" ('311 Claims 2, 3)[12]

Complainant's Proposed Construction	Respondent's Proposed Construction
"the second values"	Indefinite

[12] Although the term "the second scores" does not appear in claim 3, a construction of "the second scores" would affect claim 3 because it depends from claim 2.

The proper construction of "the second scores" is "the second values." Google's construction corrects an obvious clerical error and is consistent with a POSITA's understanding of the claim. Ex. 6 (Anderson Decl.) at ¶ 131. Sonos's only stated basis for contending this term is indefinite is lack of antecedent basis.[13] *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.

First, the '311 patent claims themselves confirm that "second scores" was intended to mean "second values." Dependent claim 2 recites: "in accordance with a determination that the first value is ***not higher*** than the second scores, ***forgoing responding*** to the detected input." Claim 2 is effectively the inverse of the last element of its independent claim 1, which recites, "in accordance with a determination that the first value is ***higher*** than the second values, ***responding*** to the detected input." '311 patent, claims 1-2 (emphasis added). A POSITA would understand from the parallel structure of the language in claims 1 and 2 that "the second scores" (of claim 2) is intended to refer to the same "second values" of independent claim 1. Ex. 6 (Anderson Decl.) at ¶ 132. Indeed, claims 1 and 2 together recite a comparison with two alternative directives for the claimed device: claim 1 covers what the device should do if the first value is higher than the second value; claim 2 adds what the device should do if the second is higher than the first. It would make no sense to compare two different ***types*** of metrics (i.e., scores with values), especially

[13] Sonos asserts in its contentions that "the second scores" is indefinite "[f]or at least the reason that the term lacks antecedent basis," Ex. 10 at 7, but Sonos has confirmed that its position regarding this term is based solely on the alleged lack of antecedent basis.

when claims 1 and 2 are directed to the same comparison to determine whether the first electronic device should respond to the detected input. This is particularly so given that the patent makes no reference to a comparison of a "score" to a "value," and a POSITA would recognize that the claim limitations recite alternative responsive steps to a single comparison (whether the first value is higher or lower). Ex. 6 (Anderson Decl.) at ¶ 133.

Claim 4, which also depends from claim 1, confirms that "the second scores" was an error. Claim 4 recites the same limitation as claim 2, referring to the correct term "the second values," rather than "scores": "The method of claim 1, further comprising: in accordance with a determination that the ***first value is not higher than the second values***, identifying a criterion associated with the voice input." A POSITA reading claims 1 and 4 would understand that "the second scores" in claim 2 was an error and that the term should be "the second values." Ex. 6 (Anderson Decl.) at ¶ 134.

Second, the prosecution history confirms that the phrase "the second scores" appears in claim 2 in error. In a preliminary amendment, the Applicant revised then-pending claim 1 to replace the term "score" with "value," and added new claims, including those that issued as claims 2 and 3. Ex. 12 ('311 File History) at 4 (Mar. 19, 2020 Prelim. Amend.) ("[Claim 1:] in accordance with a determination that the first <u>value</u> ~~score~~ is higher than the second <u>values</u> ~~scores~~, responding to the detected input . . . [Claim 2:] in accordance with a determination that the first score is not higher than the second scores, forgoing responding to the detected input.") (annotations original). After the claims were allowed,[14] the Applicant corrected claim 2 "for clarity and to correct for informalities." *Id.* at 26 (Nov. 25, 2020 Response to Office Action). In that Office Action

[14] The Office Action allowed the claims pending the Applicant's filing of a terminal disclaimer. Ex. 12 ('311 File History) at 11-16, 26.

Response, the Applicant corrected the first instance of "score" to "value," without correspondingly correcting the second instance of "scores" to "values" to be consistent with claims 1 and 4. *Id.* at 22-26; Ex. 6 (Anderson Decl.) at ¶¶ 135-37. The prosecution history thus demonstrates the Applicant's intent to correct "scores" to "values."

The only basis Sonos has articulated for its indefiniteness position is that "second scores" has no antecedent basis in the claim. Ex. 10 (Sonos Invalidity Contentions) at 7. But there can be no reasonable debate that the Applicant's failure to amend "the second scores" during prosecution (as it did with the first instance of "score" in the claim) was an error that would have been obvious to a POSITA. Ex. 6 (Anderson Decl.) at ¶¶ 131, 135-37. For the reasons set forth above, a POSITA would understand with reasonable certainty that "second scores" means "second values" in claim 2 of the '311 patent. Ex. 6 (Anderson Decl.) at ¶ 131.

H. "the detected output" ('311 Claim 12)

Complainant's Proposed Construction	Respondent's Proposed Construction
"the detected input"	Indefinite

Similar to "the second scores," a POSITA would understand "the detected output" to be an obvious typographical error. *See supra* Section VI.G. As the claims and specification show, the '311 patent is focused on detecting voice ***inputs***, not outputs, and on determining which device should respond to a given voice input. Sonos again seeks to invalidate Google's patent claims based on a clerical error, ignoring the clear intent of the patentee.[15] But the only reasonable interpretation of "the detected output" in claim 12 of the '311 patent is "the detected input."

[15] Sonos asserts in its contentions that "the detected output" is indefinite "[f]or at least the reason that the term lacks antecedent basis," Ex. 10 at 7, but Sonos has confirmed that its position regarding this term is based solely on the alleged lack of antecedent basis.

First, apart from the error in claim 12, the '311 patent claims only recite "voice input[s]"—never outputs. '311 patent, cls. 1-4, 8-11, 16, 20 (all reciting "a[/the] detected input," not "output"). Significantly, claim 10, from which at-issue claim 12 depends, recites "detecting a voice input," "determining a first value associated with the voice input," "receiving second values associated with the voice input," and then "***responding*** to the detected input." The first electronic device of claim 10 is handling the detected input, and its dependent claim 12 adds a limitation whereby the first electronic device "communicat[es] to the other devices . . . to ***forgo responding*** to the detected [input]," i.e., the input to which the first electronic device responded. These claims would not make sense together—as they must—if the first electronic device told the other devices to forgo detecting an output instead of an input. A POSITA would understand that "output" was nothing more than a clerical error. Ex. 6 (Anderson Decl.) at ¶ 140.

Second, the specification supports Google's construction. In every described instance in which a device forgoes a response, it forgoes responding to a "detected voice input," never to an output. *See* '311 patent, Fig. 6 (step 1210, "wherein the other devices of the plurality of electronic devices forgo outputting an audible response to the detected voice input;" step 1212, "forgo outputting a response to the detected voice input"), 3:8-15, 3:32-39, 3:57-64, 22:18-20 ("the leaders forgo receiving, detecting, and processing subsequent voice inputs from the user").[16] Indeed, the '311 patent uses the word "detect[ed/ing]" in connection with "voice input" more than forty times; the ***only*** instance of "detected voice output" is the error in claim 12.[17] The intrinsic

[16] Although the specification includes embodiments that forgo "outputting responses" in response to a detected input, nowhere does it disclose ***detecting*** an output. *See* '311 patent, 10:9-10, 22:4-7, 23:16-17, 29:50-52, 30:7-8, 30:20-24, 30:58-60.

[17] The limitation "detected output" was added during prosecution. Claim 12 was added during prosecution in a March 19, 2020 preliminary amendment and then allowed on the next office action, pending a terminal disclaimer. Ex. 12 ('311 File History) at 6 (Mar. 19, 2020 Prelim. Amdmt.), 15 (Nov. 5, 2020 Non-Final Office Action).

record thus supports adopting Google's proposed construction. *See Howmedica Osteonics Corp. v. DePuy Orthopaedics, Inc.*, No. 11-CV-6498 SDW MCA, 2013 WL 3455727 at *28 (D.N.J. July 9, 2013) (correcting "inner surface" to "outer surface" where "[t]hroughout the claim language and patent specification, the [patent] identifies the locking member being located on the outer part of the bearing").

Sonos contends that "the detected output" is indefinite because it lacks antecedent basis. Ex. 10 (Sonos Invalidity Contentions) at 7. For the reasons discussed above, a POSITA would understand the term in the context of the '311 patent, and Sonos's argument should be rejected.

VII. '615 PATENT

A. Overview of the '615 Patent

The invention of the '615 patent provides "more user-friendly, secure, and effective mechanisms for device commissioning than" the known prior art commissioning processes. '615 patent, 6:35-63. Specifically, the '615 patent invention simplifies the process of adding a new device to a communication network, like a wireless local area network, by using a smart wireless device, like a cellphone, tablet, or other device, that is already part of that communication network to "commission" (i.e., add) the new device onto the network. *Id.*, 7:47-55. By using a smart wireless device, a user is not required to take cumbersome steps to add new devices to their existing networks. *Id.*, 4:24-30, 6:59-65, 7:64-67, 8:9-14. Additionally, by using a short-range proximity-based transmission path, the improved commissioning process in the '615 patent is more secure than other alternatives; the wireless device must be close to the new device to carry out the commissioning process. *Id.*, 6:35-63. To do this, the '615 patent uses proximity-based transmission paths like Bluetooth communication, radio frequency identification (RFID), Zigbee, radio frequency for consumer electronics (RF4CE), Z-Wave, and near field communication. *Id.*, 8:50-54, 62-67. This improves on prior art commissioning techniques by eliminating the need for

physical and/or administrative access to an access point, special user skills, the creation of ad hoc networks, the disruption of the existing connections to the communication network, or potential compromises to sensitive data. *Id.*, 4:24-30, 6:59-65, 7:64-67, 8:9-14.

B. "short range signal transmission path" (All Claims)

Complainant's Proposed Construction	Respondent's Proposed Construction
No construction necessary (plain and ordinary meaning)	Indefinite

The '615 patent describes using a short-range signal transmission path between two devices: a "wireless device" that is part of a communication network, and a "target device" that is not yet part of that communication network.[18] Specifically, claim 1 of the '615 patent recites an inventive "method for commissioning a target device onto a communication network" that includes steps of "establishing a short-range signal transmission path with a wireless device" and "receiving one or more network configuration packets from the wireless device via the short-range signal transmission path." '615 patent, cl. 1. Dependent claim 5 specifies that this "short-range signal transmission path" can be "based on one of Bluetooth communication, radio frequency identification (RFID), Zigbee, radio frequency for consumer electronics (RF4CE), Z-Wave, and near-field communication." *Id.*, cl. 5. Claims 11 and 16 likewise describe devices storing instructions for "establishing a short-range signal transmission path" and "receiving one or more network configuration packets from the wireless device via the short-range signal transmission path." *Id.*, cls. 11, 16.

[18] *See, e.g.*, '615 patent, claim 1 ("establishing a short-range signal transmission path with a wireless device independently of the communication network"), claim 11, claim 16; *see also id*, 8:9-13 ("In one embodiment, the smart wireless device 321 utilizes a side band, proximity based, communication channel 322 that is established between the smart wireless device 321 and the other device 307 to be commissioned onto the WLAN 304.").

A POSITA would understand that a "short-range signal transmission path" in the '615 patent is a communications channel established using well-known short-range protocols identified in the '615 patent, or similar protocols. For example, the '615 patent teaches using a "proximity based communication channel" for transmissions between the "smart wireless device" that initiates the commissioning and the "other device" (i.e., target device) that is being commissioned. '615 patent, 8:9-13 ("the smart wireless device 321 utilizes a side band, proximity based, communication channel 322 that is established between the smart wireless device 321 and the other device 307 to be commissioned onto the WLAN"). This communication channel is used by the "transmitting and receiving elements" of the wireless device and the other device. *Id*., 8:10-13-14. The patent also explains that these "elements" for transmitting and receiving may use specific exemplary radio frequencies (*id*., 8:55-58), modulation techniques (*id.*, 8:58-62), and "communications standards such as, but not limited to, near field communication (NFC), RF identification (ID), BLUETOOTH LOW ENERGY, Zigbee, radio frequency for consumer electronics (RF4CE), Z-Wave, or other proprietary standards." *Id*., 8:62-67. The '615 patent claims likewise confirm these protocols are each "short-range signal" protocols. *Id*., cl. 5 (identifying Bluetooth, RFID, Zigbee, RF4CE, Z-Wave, and NFC as examples of the short-range signal transmission path in claim 1).

A POSITA would understand what a "short-range signal transmission path" is after reading the '615 patent claims, specification, and file history. As described above, multiple objective examples are provided in the specification. The phrase "short-range" is also common in the field and well understood by persons of skill. Indeed, ***Sonos itself uses "short-range" in its own***

patents, in the same way the '615 patent uses the phrase – to describe proximity-based protocols such as NFC and Bluetooth Low Energy.[19]

Sonos nonetheless argues "short-range signal" is indefinite in the '615 patent because "short-range" is a term of degree that, Sonos claims, has no objective boundary. This is wrong. As explained above, the '615 patent identifies a short-range signal as a "proximity based communication channel," and discloses (and claims) multiple examples of such "short-range" signals, including NFC, RFID, BLUETOOTH LOW ENERGY, Zigbee, RF4CE, and Z-Wave. *Id.*, 8:62-67. In 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. *Sonix Tech. Co. v. Publications Int'l, Ltd.*, 844 F.3d 1370, 1379 (Fed. Cir. 2017); *see also Niazi Licensing Corp. v. St. Jude Med. S.C., Inc.*, 30 F.4th 1339, 1349 (Fed. Cir. 2022) (terms "resilient" and "pliable" not indefinite as "examples in the written description helped provide sufficient guidance to render the claims not invalid as indefinite"). This is all that is required for definiteness. *See Sonix*, 844 F.3d at 1379 ("visually negligible" definite where specification "includes examples of visually-negligible indicators" and discloses non-numerical criteria, including "differentiability, brightness, and homogeneity," allowing one to "compare the examples and criteria from the written description . . . to determine whether an indicator is visually negligible").

C. "in response to receiving the request, receiving one or more network configuration packets" (All Claims)

Complainant's Proposed Construction	Respondent's Proposed Construction

[19] *See, e.g.*, Ex. 13 (U.S. Patent App. Pub. No. 2022/0244905), [0130]-[0131] (describing a "tap event" "generated using ***a short-range wireless protocol***" such as "BLUETOOTH LOW ENERGY (BLE)" or "near-field communication (NFC"); Ex. 14 (U.S. Patent App. Pub. No. 2022/0224971), claim 9 (claiming program instructions to stream "retrieved media in accordance with ***a short-range wireless standard***").

No construction necessary (plain and ordinary meaning)	Indefinite

The '615 patent's inventive method to "commission," or add, a new device (the "target device" in the claims) onto a communication network begins with a "request" to start that process—i.e., "a request to initiate a commissioning process" with the target device.[20] Using this method, "in response to receiving ***the request***" (i.e., the "request to initiate a commissioning process"), "configuration information" that lets the target device join the wireless network is received.[21] The '615 Patent explains the commissioning process in detail, including how commissioning requests and configuration packets are received by the target device. Some embodiments place the target device "into a commissioning mode" before the target device receives network configuration data; other embodiments let the target device receive commissioning messages right away. *Id*., 11:40-45. Some embodiments use a single commissioning message for the commissioning process; other embodiments employ multiple messages to carry out commissioning. *Id*., 11:46-48. In claims 1, 11, and 16, multiple messages are used for commissioning, and the initial request to initiate commissioning kicks off the commissioning process. *Id*., claims 1, 11, 16.

None of these steps are ambiguous or unclear. Sonos has not explained at all—much less by clear and convincing evidence—why "in response to receiving the request [to initiate a commissioning process with the target device], receiving one or more network configuration packets" would be indefinite.

[20] *See, e.g.*, '615 patent, cls. 1, 11, 16 ("receiving a request to initiate a commissioning process with the target device").

[21] *See id.*

During the meet and confer process, Sonos could not articulate why it believes this claim term is indefinite. Sonos's interrogatory responses on indefiniteness similarly fail to articulate an argument; they merely state "one of ordinary skill in the art would not understand ***how*** a device receives configuration packets in response to receiving the claimed request and the specification does not provide any guidance." Ex. 10 (Sonos Invalidity Contentions), (emphasis added). But an alleged failure to disclose "how" a claim term is performed is not a basis for indefiniteness.[22] It is thus unclear why Sonos contends this straightforward term is indefinite. If Sonos actually does articulate an indefiniteness argument, Google reserves the right to address that argument in its responsive brief.

VIII. CONCLUSION

For the reasons discussed above, Complainant respectfully requests that the ALJ adopt Complainant's proposed constructions.

Dated: December 12, 2022

Respectfully submitted,

/s/ S. Alex Lasher

S. Alex Lasher
Jeff Gerchick
Nina S. Tallon
Quinn Emanuel Urquhart & Sullivan, LLP
1300 I Street, NW, Suite 900
Washington, DC 20005
Tel.: (202) 538-8000

David A. Nelson
Quinn Emanuel Urquhart & Sullivan, LLP
191 N. Wacker Dr., Suite 2700
Chicago, IL 60606

[22] *See, e.g.*, *TQ Delta, LLC v. 2Wire, Inc.*, 373 F. Supp. 3d 509, 524 (D. Del. 2019) ("[t]he asserted claims are clear on their face as to what is claimed" and an argument that "no embodiment demonstrating ***how*** to" perform the claim term "is not an issue of whether a person of ordinary skill in the art would understand the claim scope, but an issue of whether that claim is properly supported by written description or enabled by the patent specification").

Tel.: (312) 705-7400

Patrick D. Curran
Quinn Emanuel Urquhart & Sullivan, LLP
111 Huntington Ave, Suite 520
Boston, MA 02199
Tel.: (617) 712-7100

Counsel for Complainant Google LLC

CERTIFICATE OF SERVICE

I, Jeffrey Matthews, hereby certify that on December 12, 2022, copies of the foregoing documents were served upon the following parties as indicated:

The Honorable Katherine Hiner Acting Secretary to the Commission U.S. International Trade Commission 500 E Street, S.W. Washington, D.C. 20436	☐ Via First Class Mail ☐ Via Hand Delivery ☐ Via Express Delivery ☒ Via EDIS
The Honorable Cameron R. Elliot U.S. International Trade Commission 500 E Street, S.W., Room 317 Washington, D.C. 20436 Email: michael.turner@usitc.gov	☐ Via First Class Mail ☐ Via Hand Delivery ☒ Via Electronic Mail ☐ Via EDIS
Jordan L. Coyle, Esq. ORRICK, HERRINGTON & SUTCLIFFE LLP Columbia Center 1152 15th Street, NW Washington, DC 20005 Tel.: (202) 339-8400 Email: jcoyle@orrick.com Service Email: Sonos-1330-service@orrick.com *Counsel for Respondent Sonos, Inc.*	☐ Via First Class Mail ☐ Via Hand Delivery ☐ Via Express Delivery ☒ Via Electronic Mail

/s/ Jeffrey Matthews
Jeffrey Matthews