Trials@uspto.gov
571-272-7822

Paper 7
Entered: March 11, 2021

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

AMAZON.COM, INC., AMAZON.COM LLC, AMAZON WEB SERVICES, INC., A2Z DEVELOPMENT CENTER, INC. D/B/A LAB126, RAWLES LLC, AMZN MOBILE LLC, AMZN MOBILE 2 LLC, AMAZON.COM SERVICES, INC. F/K/A AMAZON FULFILLMENT SERVICES, INC., and AMAZON.COM SERVICES LLC (FORMERLY AMAZON DIGITAL SERVICES LLC),
Petitioner,

v.

VB ASSETS, LLC,
Patent Owner.

IPR2020-01390
Patent 7,818,176

Before MICHELLE N. WORMMEESTER, SCOTT C. MOORE, and SEAN P. O'HANLON, *Administrative Patent Judges*.

O'HANLON, *Administrative Patent Judge*.

DECISION
Denying Institution of *Inter Partes* Review
35 U.S.C. § 314

I. INTRODUCTION

A. Background

Amazon.com, Inc., Amazon.com LLC, Amazon Web Services, Inc., A2Z Development Center, Inc. d/b/a Lab126, Rawles LLC, AMZN Mobile LLC, AMZN Mobile 2 LLC, Amazon.com Services, Inc. f/k/a Amazon Fulfillment Services, Inc., and Amazon.com Services LLC (formerly Amazon Digital Services LLC) (collectively, "Petitioner") filed a Petition for *inter partes* review of claims 1–52 ("the challenged claims") of U.S. Patent No. 7,818,176 B2 (Ex. 1001, "the '176 patent"). Paper 1 ("Pet."), 1. VB Assets, LLC ("Patent Owner") filed a Preliminary Response. Paper 6 ("Prelim. Resp.").

Institution of an *inter partes* review is authorized by statute only when "the information presented in the petition . . . and any response . . . shows that there is a reasonable likelihood that the petitioner would prevail with respect to at least 1 of the claims challenged in the petition." 35 U.S.C. § 314(a) (2018). We have authority, acting on the designation of the Director, to determine whether to institute an *inter partes* review under 35 U.S.C. § 314 and 37 C.F.R. § 42.4(a). For the reasons set forth below, upon considering the parties' briefs and evidence of record, we conclude that the information presented in the Petition fails to establish a reasonable likelihood that Petitioner will prevail in showing the unpatentability of any of the challenged claims. Accordingly, we decline to institute an *inter partes* review.

B. Real Parties in Interest

Petitioner identifies each of its individual entities as the real parties in interest. Pet. 2.

Patent Owner identifies itself as the sole real party in interest. Paper 4, 2.

C. Related Matters

The parties indicate that the '176 patent is the subject of the following district court proceeding:

> *VB Assets, LLC v. Amazon.com Inc.*, No. 1:19-cv-01410 (D. Del. filed July 29, 2019).

Pet. 2; Paper 4, 2. Patent Owner further notes various petitions for *inter partes* review concerning separate patents. Paper 4, 2.

D. The Challenged Patent

The '176 patent discloses a system for "selecting and presenting advertisements based on natural language processing of voice-based input." Ex. 1001, 1:8–10. Figure 3 illustrates a method of using the system and is reproduced below:

Error! Reference source not found.
Error! Reference source not found.

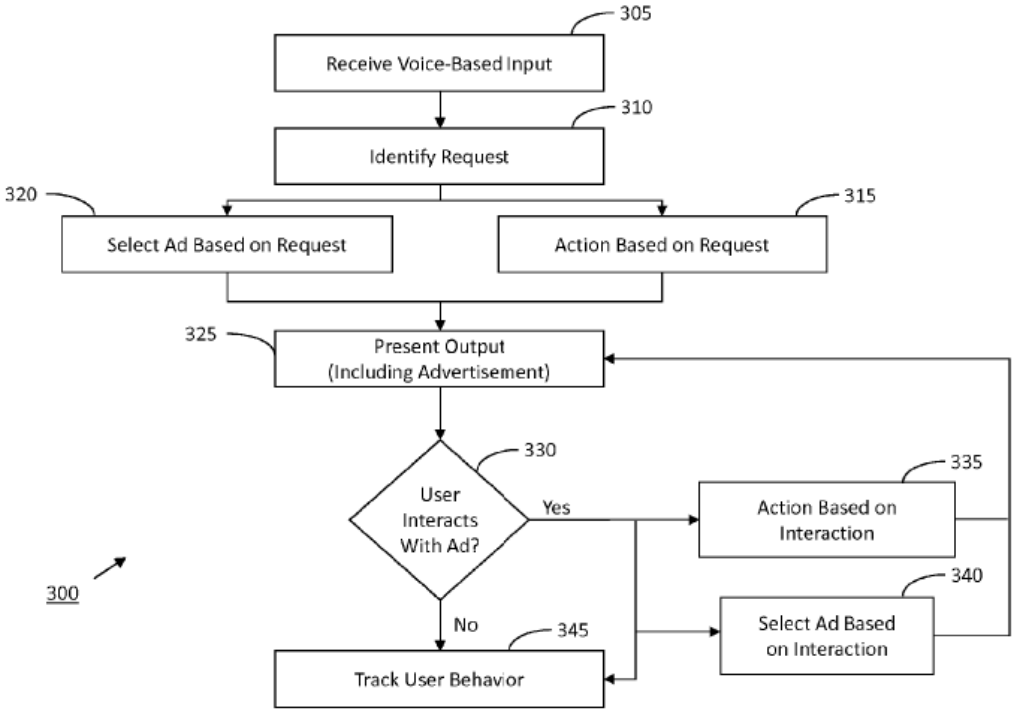

Figure 3

Figure 3 "illustrates a flow diagram of an exemplary method for selecting and presenting advertisements based on voice-based inputs." *Id.* at 2:55–57. The method begins with receiving voice-based input, also referred to as an utterance, from a user (step 305). *Id.* at 7:1–4. One or more requests within the input are then identified (step 310). *Id.* at 7:10–11. The requests can include, for example, a request for information, such as a navigation route, or to perform a task, such as placing a telephone call. *Id.* at 7:11–31. The requests may be recognized by processing the input using an automatic speech recognizer that generates one or more preliminary interpretations of the utterance using various techniques. *Id.* at 3:35–51. The requests may be part of a conversational interaction between the user and the system, whereby the interpretation can be based on previous utterances or a request can be reinterpreted based on subsequent utterances and requests. *Id.*

at 3:52–65, 7:32–48. The system performs the requested action (step 315), which may include interaction with one or more applications. *Id.* at 3:66–4:1, 7:58–66. Example applications include a navigation application, an advertising application, a music application, and an electronic commerce application. *Id.* at 4:6–9. Information in the input is also communicated to an advertising server to select one or more advertisements related to the request (step 320). *Id.* at 7:66–8:5. The advertisement and any result of the action are then presented to the user (step 325) in various manners, such as via an audible response or a display device. *Id.* at 8:6–24, 10:30–51. The advertisement may be interactive, and subsequent actions can be taken (step 335) and additional advertisements selected (step 340) based on the user's interaction with the advertisement (step 330). *Id.* at 10:52–11:10. The system can track the user's interaction with advertisements (step 345) to tailor the selection of future advertisements to the user. *Id.* at 11:11–35.

E. The Challenged Claims

Petitioner challenges claims 1–52 of the '176 patent. Pet. 1, 3–5. Claims 1, 14, 27, and 40 are independent. Claim 1 is illustrative of the challenged claims and is reproduced below:

> 1. A method for selecting and presenting advertisements in response to processing natural language utterances, comprising:
>
> receiving a natural language utterance containing at least one request at an input device;
>
> recognizing one or more words or phrases in the natural language utterance at a speech recognition engine coupled to the input device, wherein recognizing the words or phrases in the natural language utterance includes:

> mapping a stream of phonemes contained in the natural language utterance to one or more syllables that are phonemically represented in an acoustic grammar; and
>
> generating a preliminary interpretation for the natural language utterance from the one or more syllables, wherein the preliminary interpretation generated from the one or more syllables includes the recognized words or phrases;
>
> interpreting the recognized words or phrases at a conversational language processor coupled to the speech recognition engine, wherein interpreting the recognized words or phrases includes establishing a context for the natural language utterance;
>
> selecting an advertisement in the context established for the natural language utterance; and
>
> presenting the selected advertisement via an output device coupled to the conversational language processor.

Ex. 1001, 12:5–32.

F. Asserted Grounds of Unpatentability

The Petition relies on the following prior art references:

Name	Reference	Exhibit
Kennewick	US 2004/0193420 A1, published Sept. 30, 2004	1003
Yonebayashi	JP 2002-297626A, published Oct. 11, 2002	1015[1]
Jong	US 6,173,250 B1, issued Jan. 9, 2001	1018
Colledge	US 7,774,333 B2, issued Aug. 10, 2010	1019

[1] Exhibit 1015 is a certified translation (*see* Ex. 1016) of the original Japanese document (Ex. 1017).

Petitioner asserts the following grounds of unpatentability:

Claims Challenged	35 U.S.C. §	References
1–3, 6–19, 22–29, 32–45, 48–52	103(a)[2]	Kennewick, Yonebayashi, Jong
4, 5, 20, 21, 30, 31, 46, 47	103(a)	Kennewick, Yonebayashi, Jong, Colledge

Pet. 3–4. Petitioner submits a declaration of Padhraic Smyth, Ph.D. (Ex. 1002, "the Smyth Declaration") in support of its contentions.

II. ANALYSIS

A. Principles of Law

Petitioner bears the burden of persuasion to prove unpatentability, by a preponderance of the evidence, of the claims challenged in the Petition. 35 U.S.C. § 316(e). This burden never shifts to Patent Owner. *Dynamic Drinkware, LLC v. Nat'l Graphics, Inc.*, 800 F.3d 1375, 1378 (Fed. Cir. 2015). The Board may authorize an *inter partes* review if we determine that the information presented in the Petition and Patent Owner's Preliminary Response shows that there is a reasonable likelihood that Petitioner would prevail with respect to at least one of the claims challenged in the Petition. 35 U.S.C. § 314(a).

A patent claim is unpatentable under 35 U.S.C. § 103(a) if the differences between the claimed subject matter and the prior art are such that

[2] The application resulting in the '176 patent was filed on a date prior to the date when the Leahy-Smith America Invents Act ("AIA"), Pub. L. No. 112–29, 125 Stat. 284 (2011), took effect. Thus, we refer to the pre-AIA version of section 103.

the subject matter, as a whole, would have been obvious at the time the invention was made to a person having ordinary skill in the art to which the subject matter pertains. *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). The question of obviousness is resolved on the basis of underlying factual determinations including: (1) the scope and content of the prior art; (2) any differences between the claimed subject matter and the prior art; (3) the level of skill in the art; and (4) when in evidence, any objective evidence of non-obviousness.[3] *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

B. Level of Ordinary Skill in the Art

Petitioner contends that a person having ordinary skill in the art at the time of the invention "would have at least a Bachelor-level degree in computer science, computer engineering, electrical engineering, or a related field in computing technology, and two years of experience with automatic speech recognition and natural language understanding, or equivalent education, research experience, or knowledge." Pet. 4.

Patent Owner does not contest Petitioner's definition or proffer an alternate definition. *See generally* Prelim. Resp.

The level of ordinary skill in the art often is evidenced by the references themselves. *See Okajima v. Bourdeau*, 261 F.3d 1350, 1355 (Fed. Cir. 2001); *In re GPAC Inc.*, 57 F.3d 1573, 1579 (Fed. Cir. 1995); *In re Oelrich*, 579 F.2d 86, 91 (CCPA 1978). The level of ordinary skill proposed by Petitioner appears to be consistent with that of the references,

[3] At this stage of the proceeding, the parties have not directed us to any such objective evidence.

and we apply Petitioner's proposed level of ordinary skill for purposes of this Decision.

C. Claim Construction

In an *inter partes* review, claims are construed using the same claim construction standard that would be used to construe the claims in a civil action under 35 U.S.C. § 282(b), including construing the claims in accordance with the ordinary and customary meaning of such claims as understood by one of ordinary skill in the art and the prosecution history pertaining to the patent. 37 C.F.R. § 42.100(b) (2019). Thus, we apply the claim construction standard as set forth in *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005) (en banc). In addition to the specification and prosecution history, we also consider use of the terms in other claims and extrinsic evidence including expert and inventor testimony, dictionaries, and learned treatises, although extrinsic evidence is less significant than the intrinsic record. *Id.* at 1312–17. Usually, the specification is dispositive, and it is the single best guide to the meaning of a disputed term. *Id.* at 1315.

Only those terms that are in controversy need be construed, and only to the extent necessary to resolve the controversy. *Nidec Motor Corp. v. Zhongshan Broad Ocean Motor Co.*, 868 F.3d 1013, 1017 (Fed. Cir. 2017) (citing *Vivid Techs., Inc. v. Am. Sci. & Eng'g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999)).

Petitioner asserts that "[t]he challenged claims should be interpreted in accordance with [37 C.F.R.] § 42.100(b)." Pet. 8.

Patent Owner notes that our rules require a petition to set forth how the challenged claims are to be construed and that "the [P]etition does not

explicitly construe any terms." Prelim. Resp. 3. Patent Owner argues that construction of "acoustic grammar" as recited in claim 1 is necessary to understanding Petitioner's arguments. *Id.* Patent Owner argues that "Petitioner[] also telegraph[s] claim construction gamesmanship" and, therefore, that we should deny institution. *Id.* at 4.

We are not persuaded that we should exercise our discretion to deny institution based on Petitioner's alleged failure to set forth adequate claim constructions. By arguing that the claim terms should be construed according to their "ordinary and customary meaning" (Pet. 7–8), Petitioner has complied with our rule that the Petition must identify how the challenged claims are to be construed. *See* 37 C.F.R. § 42.104(b)(3). Nor do Patent Owner's assertions of a possibility of "gamesmanship" provide a reason compelling us to deny institution, as Patent Owner's arguments are merely speculation about possible future actions.

However, we agree that we must interpret the term "acoustic grammar." "Acoustic grammar" does not appear in the Specification of the '176 patent. *See generally* Ex. 1001. This term was added to the claims via amendment on February 17, 2010. Ex. 1008, 257–69. The Applicant added claim 22, which contained the "mapping" and "generating" recitations of challenged claim 1 (*id.* at 263), and claim 26, which contained the "map" and "generate" recitations of challenged claim 27 (*id.* at 264). The Examiner indicated that these added claims "would be allowable if rewritten in independent form" because "the prior art of record does not disclose mapping a stream of phonemes to one or more syllables that are phonemically represented in an acoustic grammar and generating a preliminary interpretation from the one or more syllables (see related U.S.

Patent 7,634,409[4] assigned to the instant application's assignee)." *Id.* at 304–05. The Applicant subsequently amended claims 22 and 26 to be in independent form. *Id.* at 327–29. The Examiner then allowed the claims, noting that "new claims 29-54 find support in the [S]pecification, either directly or through an incorporated by reference application." *Id.* at 354. Claim 22 issued as challenged claim 1 and claim 26 issued as challenged claim 27. *Id.* at 356.

Kennewick '409 discloses that "the performance of the speech engine may be improved by using phoneme recognition." Ex. 1020, 2:38–40. "Phonemes are distinct units of sound. For example, the word 'those' is made up of three phonemes; the first is the 'th' sound, the second is the 'o' sound, and the third is the 's' sound." Ex. 1008, 61 (WO 01/78065 A1, page 2). "Each phoneme has distinguishable acoustic characteristics and, in combination with other phonemes, forms larger units such as syllables and words." Ex. 1011, 22; *see also id.* at 64 (presenting a list of 42 phonemes for the English language). "Phoneme recognition may be based on any suitable acoustic grammar that maps a speech signal into a phonemic representation." Ex. 1020, 2:46–48. "Characteristics of a speech signal may be mapped to a phonemic representation to construct a suitable acoustic grammar . . . ." *Id.* at 6:16–18.

> For example, the English language may be mapped into a detailed acoustic grammar representing the phonotactic rules of English, where words may be divided into syllables, which may further be divided into core components of an onset, a nucleus,

[4] We note that this patent is incorporated into the '176 patent (Ex. 1001, 3:46–51) and is included in the record as Exhibit 1020 ("Kennewick '409").

> and a coda, which may be further broken down into one or more sub-categories.

Id. at 6:21–26. "[A] real-world acoustic grammar modeled after a language is likely to have a maximum of roughly fifty phonemes." *Id.* at 7:24–26.

"[A]coustic grammars may be formed as trees with various branches representing many different syllables forming a speech signal." Ex. 1020, 2:53–56.

> Using the English language as an example, the grammar tree may include various branches representing English language syllables. The speech engine may traverse one or more grammar trees to generate one or more preliminary interpretations of a phoneme stream as a series of syllables that map to a word or phrase.

Id. at 6:32–38. Nodes in the grammar tree may represent words or items in a list. *Id.* at 6:53–56.

Thus, Kennewick '409 explains that an "acoustic grammar" is a collection of the phonemes, or distinct units of sound of a spoken language, linked together to form syllables, which are linked together to form the words of the language. On this record and for the purposes of this Decision, we interpret "acoustic grammar" as used in the '176 patent in the same manner.

This interpretation is consistent with use of the term in the claims. For example, claim 1 recites "mapping a stream of phonemes contained in the natural language utterance to one or more syllables that are phonemically represented in an acoustic grammar." Ex. 1001, 12:15–17. Thus, the claim requires the acoustic grammar to link the phonemes in the user's utterance to syllables, in the same manner as discussed in Kennewick '409.

Error! Reference source not found.
Error! Reference source not found.

D. Overview of the Asserted Prior Art

1. Kennewick

Kennewick discloses a system that performs "retrieval of online information and processing of commands through a speech interface in a vehicle environment." Ex. 1003 ¶ 2. Figure 5 illustrates the system and is reproduced below:

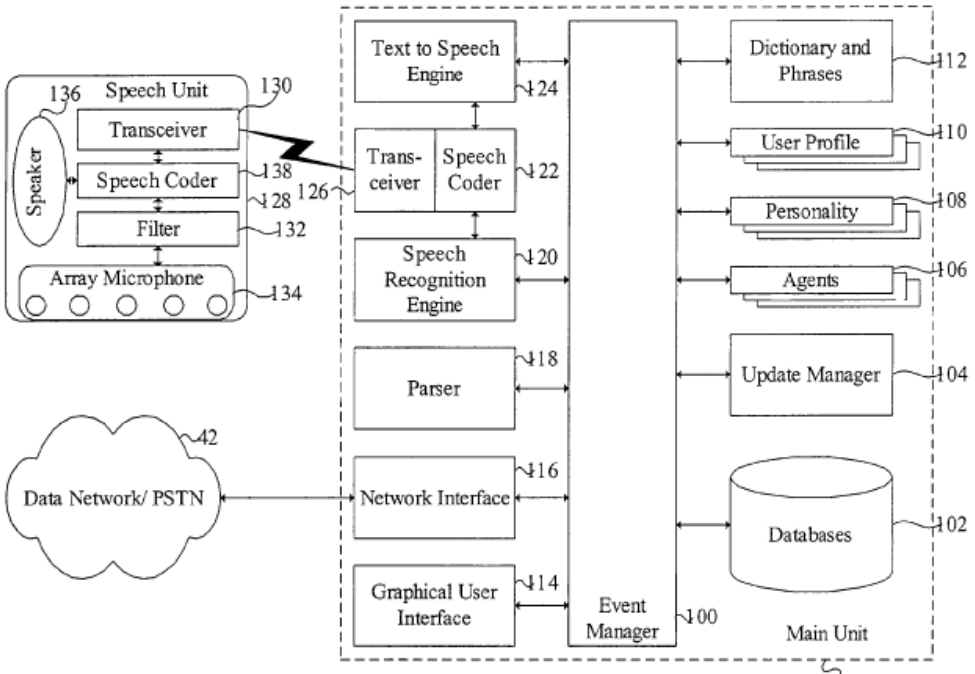

Figure 5. Speech Processing System Block Diagram

Figure 5 "shows an overall diagrammatic view of the interactive natural language speech processing system according to one embodiment of the invention." *Id.* ¶ 118. Speech unit 128 detects speech using microphone 134. *Id.* ¶ 121. The detected speech passes through filter 132 to coder 138 for encoding and compression. *Id.* The coded speech is then transmitted via transceiver 130 to transceiver 126 of main unit 98, and then decoded and decompressed by speech coder 122. *Id.* ¶¶ 121, 123, Fig. 5. Speech recognition unit 120 processes the decoded speech to detect words

and phrases. *Id.* ¶ 123. Parser 118 transforms the recognized words and phrases into complete commands and questions using data supplied by domain agents 106. *Id.* ¶¶ 123, 160. The parser determines the context for the speech, and from the context determines the domain and, thereby, the domain agent to be invoked. *Id.* ¶ 160. The agents then process the commands or questions using one or more devices under their control, and return appropriate responses to the user. *Id.* ¶¶ 120, 123–124. Generally, the agents are specific to a single domain. *Id.* ¶¶ 17, 126. In one embodiment, the system provides offers and promotions for goods and services based on the user's location. *Id.* ¶¶ 65–66.

2. Yonebayashi

Yonebayashi recognizes that various electronic devices, such as personal computers and microwave ovens, can be connected to and receive advertisement information from other devices via networks, such as a Local Area Network. Ex. 1015 ¶ 2. Yonebayashi purports to improve upon such systems by providing advertisements having more appropriate content for the user and enabling interaction between the user and the advertisement. *Id.* ¶ 8.

Yonebayashi discloses a computer-based advertisement presentation device. Ex. 1015 ¶ 17. The device includes a dictionary storage unit that stores various types of advertisement information, including active advertisement information and response advertisement information. *Id.* ¶¶ 34–36. Active advertisement information is information for the system to actively present advertisements to the user. *Id.* ¶ 35. Response advertisement information is information for advertising in accordance with the user's remarks and inquiries. *Id.* ¶ 36. The dictionary storage unit also

includes a case dictionary that contains a series of if-then rules, called cases, and user information by which the advertisement information is selected and formatted for presentation to the user. *Id.* ¶¶ 39–41. Figure 5 illustrates an example of a dialog between the system, referred to as the "agent," and a user and is reproduced below:

[FIGURE 5]

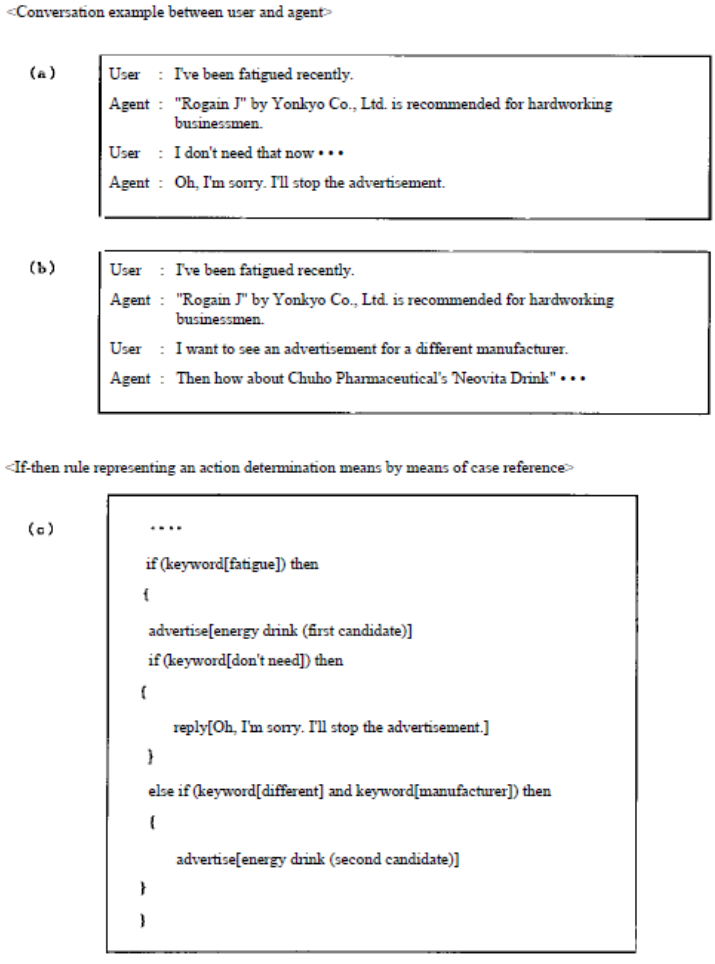

Figure 5 shows conversation examples between a user and an agent in a case of presenting an advertisement for an energy drink. *Id.* at 26. The process begins with the user saying "I've been fatigued lately." *Id.* ¶ 49. The system's character string acquisition means receives the user's remarks and the preprocessing means identifies the words therein. *Id.* ¶¶ 49–50. The

action determination unit compares the detected words to the "if" portion of the rules and, recognizing the keyword "fatigue," determines "energy drink (first candidate)" to be the appropriate action. *Id.* ¶ 50. The presentation means then presents the advertisement for the first candidate energy drink in the advertisement information dictionary. *Id.* ¶ 51. The system then awaits further user remarks and reacts appropriately. *Id.* ¶¶ 52, 53. For example, if the user indicates that no advertisements are desired, the advertisement is terminated, and if the user indicates that another brand of energy drink is desired, the system presents an advertisement for the next highest rank candidate. *Id.*

3. Jong

Jong discloses "an apparatus and method for providing real time communication over a data network." Ex. 1018, 1:8–9. Jong recognizes that known voice telephony systems that digitize voice input signals for transmission experience significant delay and distortion and require large bandwidth. *Id.* at 1:25–41. Jong purports to improve upon such systems by converting speech input signals into text data and transmitting the text data over a data network. *Id.* at 1:55–60, 3:14–17. The receiving party can display the speech input as text, and the text can also be converted into synthesized speech and audibly presented to the receiving party. *Id.* at 5:25–30.

The voice input is converted into text by speech recognition device 203. Ex. 1018, 5:14–15. Speech recognition device 203 includes spectral analysis device 301, word-level matching device 302, word model device 303, subword models database 304, and lexicon database 305. *Id.* at 5:35–40, Fig. 3.

> When the speech input signals are received by the speech recognition device 203, the spectral analysis device 301 receives the speech input signals and extracts feature vectors from them. The feature vectors are input to a word-level matching device 302[,] which compares the feature vectors against the word models retrieved by the word model device 303 to identify the words that make up the speech input signals.

Id. at 5:42–49.

> The word model device 303 includes a listing of phonemes (speech sounds)[,] which are used to identify the words in the speech input signals. The subword model database 304 contains word syllables that are correlated with the phonemes of the word model device 303. The lexicon database 305 stores a dictionary of recognizable words.

Id. at 5:51–57. "The word model device 303 identifies the phonemes in the speech input signals and extracts the corresponding syllables from the subword model database 304." *Id.* at 5:59–61. "[T]he syllables that make up the various words in the speech input signals are grouped into the recognizable words identified using the lexicon database 305." *Id.* at 5:64–67.

4. Colledge

Colledge discloses a system for associating a search query or information with an advertisement. Ex. 1019, 4:23–25. Colledge recognizes that Internet search engines perform searches based on keywords and generate revenue by selling keywords to advertisers. *Id.* at 1:25–35. Colledge further recognizes that typical systems can result in the advertiser's promotions being associated with irrelevant searches if a keyword has multiple meanings and being omitted from relevant searches if the user's keywords are not the exact same as the purchased keywords. *Id.* at 1:36–42.

Colledge purports to improve upon such systems by disambiguating the search query by identifying the intended meaning of each word in the query. Ex. 1019, 9:49–54. The system then expands the relevant search terms to include semantically related senses. *Id.* at 9:55–60. For example, for a search including the keywords "Java" and "holiday," the system can disambiguate "Java" to mean the island rather than the object-oriented programming language and can expand "holiday" to include "vacation." *Id.* at 10:34–65. The results as well as any relevant advertisements are then presented to the user. *Id.* at 10:16–20.

E. Asserted Obviousness in View of Kennewick, Yonebayashi, and Jong

Petitioner argues that claims 1–3, 6–19, 22–29, 32–45, and 48–52 would have been obvious over the combination of Kennewick, Yonebayashi, and Jong. Pet. 16–64. In support of its showing, Petitioner relies upon the Smyth Declaration. *Id.* (citing Ex. 1002). We have reviewed Petitioner's assertions and supporting evidence. For the reasons discussed below, and based on the record before us, we determine that Petitioner does not demonstrate a reasonable likelihood of prevailing in showing that claims 1–3, 6–19, 22–29, 32–45, and 48–52 would have been obvious over the combination of Kennewick, Yonebayashi, and Jong.

1. Claims 1–3, 6–13, 27–29, and 32–39

Independent claim 1 recites, in relevant part, "mapping a stream of phonemes contained in the natural language utterance to one or more syllables that are phonemically represented in an acoustic grammar." Ex. 1001, 12:15–17. Petitioner notes that "acoustic grammar" does not appear in the Specification of the '176 patent. Pet. 25. Petitioner notes that

the '176 patent incorporates Kennewick '409, which, Petitioner argues, "recognizes '[p]honeme recognition may be based on any suitable acoustic grammar that maps a speech signal into a phonemic representation'" and "'[p]ortions of a word may be represented by a syllable' in the grammar." *Id.* (alterations in original) (emphasis omitted) (quoting Ex. 1020, 2:46–51).

Petitioner relies on Jong to teach recognizing and mapping phonemes to syllables in an acoustic grammar. Pet. 25–26. Specifically, Petitioner relies on Jong's speech recognition device 203, noting that, when the speech recognition device receives speech input signals, its spectral analysis device extracts feature vectors that are input into its word-level matching device 302, which compares the feature vectors against the word models retrieved by word model device 303. *Id.* at 25 (citing Ex. 1018, 5:41–47). Petitioner notes that the word model device includes a listing of phonemes, subword model database 304 contains word syllables that are correlated with the phonemes of the word model device, and lexicon database 305 stores a dictionary of recognizable words. *Id.* at 25–26 (citing Ex. 1018, 5:51–58). "The word model device 303 identifies the phonemes in the speech input signals and extracts the corresponding syllables from the subword model database 304." *Id.* at 26 (quoting Ex. 1018, 5:59–61). Thus, Petitioner argues, "spectral analysis device 301, word-level matching device 302, and the word model device 303 are an acoustic grammar." *Id.* (citing Ex. 1002 ¶ 101: Ex. 1020, 2:46–51).

Patent Owner argues that "Petitioner[] ha[s] not cited to anything in Jong that indicates its mapping of phonemes to syllables involves an acoustic grammar" and that "it is entirely unclear what definition of acoustic grammar Petitioner[] [is] using." Prelim. Resp. 20. Patent Owner notes that

Kennewick '409 teaches that an acoustic grammar "may include 'phonotactic rules of the English language,'" "may be formed as trees with various branches representing many different syllables forming a speech signal," and "'may be represented entirely by a loop of phonemes' which 'may include a linking element between transitions.'" *Id.* at 21 (quoting Ex. 1020, 2:48–56, 2:61–67). Patent Owner interprets the Petition as defining "an acoustic grammar [to be] anything that maps phonemes to syllables," and argues that this interpretation cannot "be reconciled with the fact that such a function is already recited elsewhere in the claims." *Id.* Patent Owner notes that the Petition maps "acoustic grammar" to devices disclosed by Jong and argues that "Petitioner[] ha[s] not taken the necessary step of explaining how the devices in Jong . . . are an 'acoustic grammar.'" *Id.* at 22. "[A]sserting that the devices perform the same function as an acoustic [grammar] does not explain how they *are* an acoustic grammar." *Id.*

We agree that Petitioner's failure to advance a construction for "acoustic grammar" makes consideration of Petitioner's arguments difficult. Petitioner's citation to two sentences in Kennewick '409 also fails to provide an explanation for how Petitioner interprets the term—this is especially true given that the second citation discusses the English language rather than an acoustic grammar. *See* Ex. 1020, 2:48–53 ("For example, the English language may be broken down into a detailed grammar of the phonotactic rules of the English language. Portions of a word may be represented by a syllable, which may be further broken down into core components of an onset, a nucleus, and a coda, which may be further broken down into sub-categories."); *see also* Pet. 25 (citing Ex. 1020, 2:48–51). Notably,

Petitioner did not reference the following sentence, which states, "Various different acoustic grammars may be formed as trees with various branches representing many different syllables forming a speech signal." *Id.* at 2:53–56.

Petitioner fails to explain with requisite particularity how Jong teaches the use of an acoustic grammar. *See* 35 U.S.C. § 312(a)(3). Petitioner relies on a single sentence from the summary section of Kennewick '409 to the exclusion of the rest of the discussion regarding acoustic grammars. Pet. 25. Petitioner then summarizes functions performed by certain of Jong's devices and concludes by asserting that the devices themselves are an acoustic grammar. *Id.* at 25–26. Petitioner does not, however, discuss the structure of the components and explain how that structure is an acoustic grammar. Even if we were to agree that the functions noted by Petitioner indicate that Jong's devices *employ* an acoustic grammar, Petitioner does not explain adequately how the *devices themselves* are an acoustic grammar. *See id.* For the same reasons, Petitioner's assertion that Jong's devices are an acoustic grammar is inconsistent with our interpretation of the term set forth in § II.C above.

Moreover, Petitioner does not explain adequately why a person having ordinary skill in the art would modify Kennewick's system to include the asserted acoustic grammar of Jong. "[R]ejections on obviousness grounds cannot be sustained by mere conclusory statements; instead, there must be some articulated reasoning with some rational underpinning to support the legal conclusion of obviousness." *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006) (cited with approval in *KSR*, 550 U.S. at 418). Additionally, "one must have a motivation to combine accompanied by a reasonable

expectation of achieving what is claimed in the patent-at-issue." *Intelligent Bio-Sys, Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359, 1367 (Fed. Cir. 2016).

Petitioner asserts that an ordinarily skilled artisan would "include an acoustic grammar that recognizes a stream of phonemes in the user's speech signal and maps the recognized stream of phonemes to syllables." Pet. 26 (citing Ex. 1002 ¶ 103; Ex. 1018, 5:41–61). Describing the function performed by the asserted acoustic grammar, however, does not explain why a skilled artisan would find benefit in modifying Kennewick's device or that the artisan would have a reasonable expectation that the modification would succeed. Petitioner's declarant similarly discusses how the asserted acoustic grammar could be used without explaining the benefit of such use. Ex. 1002 ¶ 103. Petitioner also states that incorporating the asserted acoustic grammar into Kennewick's system "would enhance how words and phrases are recognized from a user's voice input." Pet. 18. Again, Petitioner does not explain adequately how Kennewick's system would be enhanced or why there would be a reasonable expectation that the combination would succeed.

For the reasons explained above, the arguments and evidence presented in the Petition fail to provide sufficient support for Petitioner's contentions that Jong discloses an acoustic grammar or that it would have been obvious to include the asserted acoustic grammar in Kennewick's system as required by claim 1 and incorporated into dependent claims 2, 3, and 6–13. Therefore, the Petition has not demonstrated a reasonable likelihood of success in challenging these claims.

Independent claim 27 recites a system that includes the components and performs the steps recited in claim 1. Ex. 1001, 15:9–38. Petitioner argues that claim 27 would have been obvious for the reasons set forth regarding claim 1. Pet. 57–59 (referencing arguments advanced for claim 1). For the reasons set forth above in our analysis of claim 1, the Petition has not demonstrated a reasonable likelihood of success in challenging claim 27 or its dependent claims 28, 29, and 32–39.

2. Claims 14–19, 22–26, 40–45, and 48–52

Independent claim 14 recites, in relevant part, "determining that the conversational language processor incorrectly interpreted the words or phrases in response to an adaptive misrecognition engine detecting a predetermined event, wherein the conversational language processor reinterprets the words or phrases in response to the predetermined event." Ex. 1001, 13:58–63. Petitioner relies on its showing for claim 12 for this recitation. Pet. 55–56. Regarding claim 12, Petitioner argues that Kennewick's system "uses a scoring system to determine whether the natural language utterance includes incomplete or ambiguous information" and, if the score is not high enough, "the system can ask a question of the user to verify the question or command is correctly understood" and "the system can ask one or more questions to attempt to resolve the ambiguity." Pet. 53–54 (citing Ex. 1003 ¶ 161). Petitioner argues that "[t]his functionality of parser 118 and agents 106 identifying a confidence score and generating a verification question or command is the adaptive misrecognition engine." *Id.* at 54. Petitioner argues that "the user may respond to the verification question or command" and "the user's subsequent utterance is the predetermined event." *Id.* According to Petitioner, "[i]f the

user's subsequent utterance provides additional information, the information is used to reinterpret the words of the user's initial utterance." *Id.* at 55.

Patent Owner argues that "the [P]etition does not identify a determination of an incorrect interpretation" because the Petition relies on Kennewick's disclosure of "a scoring system to determine whether the natural language utterance includes incomplete or ambiguous information," which, Patent Owner asserts, "is not a determination that the system made an incorrect interpretation." Prelim. Resp. 24 (quoting Pet. 53–54). Patent Owner argues that the Petition's reliance on Kennewick to disclose reinterpreting the uttered words or phrases fails because, "[r]ather than reinterpreting the original utterance, Kennewick discloses that the user simply asks the question again with additional information." *Id.* at 25. According to Patent Owner, "the [P]etition conflates Kennewick's recognition of ambiguity with a misinterpretation." *Id.*

We agree that the Petition does not explain adequately how Kennewick discloses determining its system incorrectly interpreted words or phrases in a user's utterance or reinterpreting such incorrectly interpreted words or phrases. Petitioner equates Kennewick's determination that an utterance is "incomplete or ambiguous" with a determination that it made an incorrect interpretation and subsequent reliance upon additional information with a reinterpretation of the utterance. Pet. 53–55 (citing Ex. 1003 ¶ 161). Kennewick discloses that parser 118 interprets recognized words into complete commands and questions using a scoring system to determine the most likely context for a user's question or command. Ex. 1003 ¶¶ 123, 160–161. "If the confidence level of the score is not high enough to ensure a reliable response, the system can ask a question of the user to verify the

question or command is correctly understood." *Id.* ¶ 161. If the user's response does not clarify the user's question or command, "the system can ask one or more questions to attempt to resolve the ambiguity." *Id.* "Once the context for the question or command has been determined, the parser 118 can invoke the correct agent 156, 150." *Id.* ¶ 162. Thus, Kennewick discloses that its system clarifies ambiguities in the utterance as part of interpreting the utterance and then, after the ambiguities have been clarified, acting on the user's question or command. Thus, Petitioner's reliance on Kennewick's clarification process fails to explain adequately how Kennewick first incorrectly interprets the utterance and then determines that the utterance was incorrectly interpreted.

For the reasons explained above, the arguments and evidence presented in the Petition fail to provide sufficient support for Petitioner's contentions that Kennewick determines that its system incorrectly interpreted the words or phrases in a user's utterance as required by claim 14 and incorporated into dependent claims 15–19 and 22–26. Therefore, the Petition has not demonstrated a reasonable likelihood of success in challenging these claims.

Independent claim 40 recites a system that includes the components and performs the steps recited in claim 14. Ex. 1001, 16:49–17:6. Petitioner argues that claim 40 would have been obvious for the reasons set forth regarding claim 14. Pet. 62–63 (referencing arguments advanced for claim 14). For the reasons set forth above in our analysis of claim 14, the Petition has not demonstrated a reasonable likelihood of success in challenging claim 40 or its dependent claims 41–45 and 48–52.

F. Asserted Obviousness in View of Kennewick, Yonebayashi, Jong, and Colledge

Petitioner argues that claims 4, 5, 20, 21, 30, 31, 46, and 47 would have been obvious over the combination of Kennewick, Yonebayashi, Jong, and Colledge. Pet. 64–71. Petitioner relies on Kennewick, Yonebayashi, and Jong as set forth in § II.E above, and relies on Colledge to teach updating a personal cognitive model by building statistical profiles. *Id.*

Claims 4 and 5 depend indirectly from claim 1, claims 20 and 21 depend indirectly from claim 14, claims 30 and 31 depend indirectly from claim 27, and claims 46 and 47 depend indirectly from claim 40. As Petitioner has failed to establish that Kennewick, Yonebayashi, and Jong satisfy the recitations of parent claims 1, 14, 27, and 40 as discussed above, Petitioner has likewise failed to establish how these recitations are satisfied with respect to these dependent claims. Thus, the Petition has not demonstrated a reasonable likelihood of prevailing in showing that claims 4, 5, 20, 21, 30, 31, 46, and 47 would have been obvious over the combination of Kennewick, Yonebayashi, Jong, and Colledge.

III. CONCLUSION

For the foregoing reasons, we are not persuaded that the Petition establishes a reasonable likelihood that Petitioner would prevail in any of its challenges to claims 1–52 of the ’176 patent.

IV. ORDER

In consideration of the foregoing, it is hereby ordered that the Petition is *denied*, and no trial is instituted.

Error! Reference source not found.
Error! Reference source not found.

For PETITIONER:

J. David Hadden
Saina Shamilov
Allen Wang
FENWICK & WEST LLP
dhadden-ptab@fenwick.com
sshamilov-ptab@fenwick.com
allen.wang@fenwick.com
VBAssets-IPR@fenwick.com

For PATENT OWNER:

Matthew A. Argenti
Michael T. Rosato
James C. Yoon
Ryan Smith
WILSON SONSINI GOODRICH & ROSATI
margenti@wsgr.com
mrosato@wsgr.com
jyoon@wsgr.com
rsmith@wsgr.com