

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

MICROSOFT CORPORATION,
Petitioner,

v.

DIALECT, LLC,
Patent Owner.

Case IPR2025-00658
Patent 8,195,468

PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	THE ’468 PATENT	2
III.	THE BOARD NEED NOT REACH ANY CLAIM CONSTRUCTION ISSUES TO DENY THE PETITION.....	7
IV.	PETITIONER’S CHALLENGES FAIL ON THE MERITS	8
V.	CONCLUSION.....	13

TABLE OF AUTHORITIES

Rules

37 C.F.R. § 42.107	1
37 C.F.R. § 42.120(a).....	1

PATENT OWNER’S EXHIBIT LIST

Exhibit No.	Description
2001	First Amended Docket Control Order, <i>Dialect, LLC v. Bank of America, N.A.</i> , Civ. No. 2:24-cv-207, Dkt. 52 (E.D. Tex. Aug. 28, 2024)
2002	Comparison of Districts Within the Fifth Circuit — 12-Month Period Ending September 30, 2024 (available at https://www.uscourts.gov/sites/default/files/2025-02/fcms_na_distcomparison1231.2024.pdf)
2003	Defendant’s Invalidity Contentions, <i>Dialect, LLC v. Bank of America, N.A.</i> , Civ. No. 2:24-cv-207 (E.D. Tex. Sept. 12, 2024)
2004-2006	Reserved
2007	U.S. Pat. 7,137,126 (PCT counterpart to Coffman)
2008-2011	Reserved
2012	Ex. B (Maes) to Defendant’s Invalidity Contentions, <i>Dialect, LLC v. Bank of America, N.A.</i> , Civ. No. 2:24-cv-207 (E.D. Tex. Sept. 12, 2024)
2013	Ex. B (Coffman) to Defendant’s Invalidity Contentions, <i>Dialect, LLC v. Bank of America, N.A.</i> , Civ. No. 2:24-cv-207 (E.D. Tex. Sept. 12, 2024)
2014-2018	Reserved
2019	Return of Summons, <i>Dialect, LLC v. Bank of America, N.A.</i> , Civ. No. 2:24-cv-207 (E.D. Tex. Mar. 27, 2024)
2020	“IPBC Global 2025: Acting USPTO Director says IPR use needs to change” (June 10, 2025) (available at https://www.iam-media.com/article/ipbc-global-2025-acting-uspto-director-says-ipr-use-needs-change)

Pursuant to 37 C.F.R. § 42.107, Patent Owner Dialect, LLC (“Dialect” or “Patent Owner”) files this preliminary response to the Petition, setting forth reasons why the Petition for *inter partes* review (“IPR”) of claims 19, 25-30, and 32 of U.S. Patent No. 8,195,468 (the “’468 patent”), as requested by Microsoft Corporation (“Petitioner”), should be denied.¹

I. INTRODUCTION

Petitioner’s challenges to the ’468 patent should be denied because Petitioner fundamentally misunderstands the prior art and fails to address the claimed “cognitive model” requirements. The ’468 patent claims methods that identify context stack entries matching a merged transcription and determine the most likely context for multimodal input, using semantic knowledge-based models that include both “personalized cognitive models” and “general cognitive models.” Petitioner’s invalidity challenge fails on multiple fronts to disclose these features.

First, Petitioner confuses speech-to-text conversion for post-transcription analysis by incorrectly arguing that Maes teaches the claimed context determination process. Maes actually describes phoneme scoring to determine what a user has said during speech recognition—not whether a completed utterance matches entries in a context stack as required by the claims.

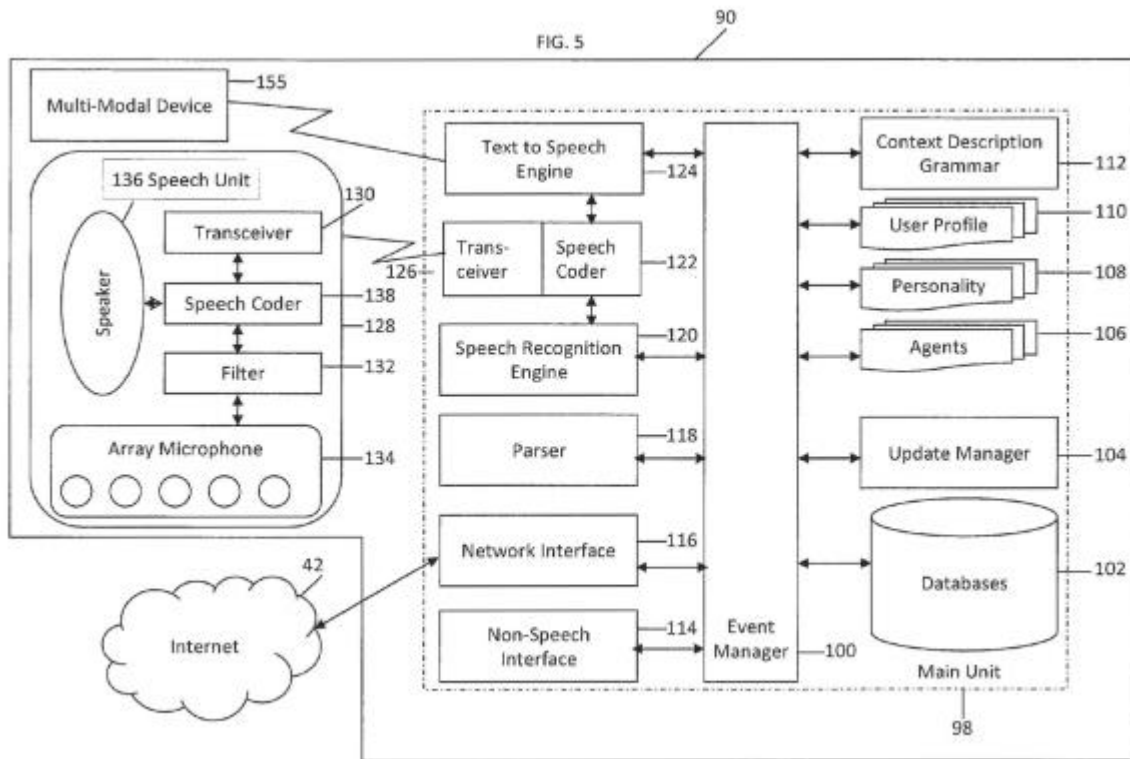
¹ By submitting this Preliminary Response, no waiver of any argument is intended by Patent Owner. Patent Owner will have a right to file a complete response to the Petition, should the Board institute *inter partes* review. 37 C.F.R. § 42.120(a).

Second, and more troubling, to support Petitioner's deficient analysis of "cognitive models," Petitioner's expert improperly altered a key quotation from the '468 patent specification, substituting "and" for "thus," resulting in a changed meaning. Rather than addressing what makes a model "cognitive"—such as the specification's requirement that cognitive models "predict a user's next action"—Petitioner's expert focused solely on whether models "assist" with speech recognition, ignoring the requirement for cognitive functionality entirely. This pattern of deficient analysis extends across all the asserted references, where the expert repeatedly fails to explain how cache databases or other disclosed features satisfy the "cognitive" requirements. Because Petitioner cannot establish that the prior art teaches the claimed cognitive model functionality, the challenges also fail and warrant denial of institution.

II. THE '468 PATENT

The '468 patent is entitled, "Mobile Systems and Methods of Supporting Natural Language Human-Machine Interactions," and relates to enabling natural language human-machine interactions with multimodal inputs (such as from a speech interface and non-speech interfaces).

FIG. 5 of the '468 patent, reproduced below, is an overall diagrammatic view of an embodiment of the '468 patent:



'468 patent, FIG. 5; *see also id.*, 15:38-40. As shown in the figure above, the system includes, among other things, speech unit 128, speech recognition engine 120, context description grammar module 112, parser 118, agents 106, multi-modal device 155 and non-speech interface 114. '468 patent, 23:36-47. Speech unit 128 includes a microphone to receive a spoken utterance from a user. '468 patent, 25:48-59. After initial processing, the speech is passed to speech recognition engine 120 for processing "using the context description grammar module 112." '468 patent, 26:18-23. Then, "[a]ny recognized information may be processed by the parser 118, which transforms information into complete algorithms and questions using data supplied by knowledge agents." *Id.* "The

knowledge agents may then process the commands or questions.” ’468 patent, 26:27-31. Non-speech interface 114 and multi-modal devices 155 “may be used as a substitute or complement to the speech interface.” ’468 patent, 26:51-53.

FIG. 8 of the ’468 patent, reproduced below, illustrates aspects of an embodiment of the ’468 patent in more detail.

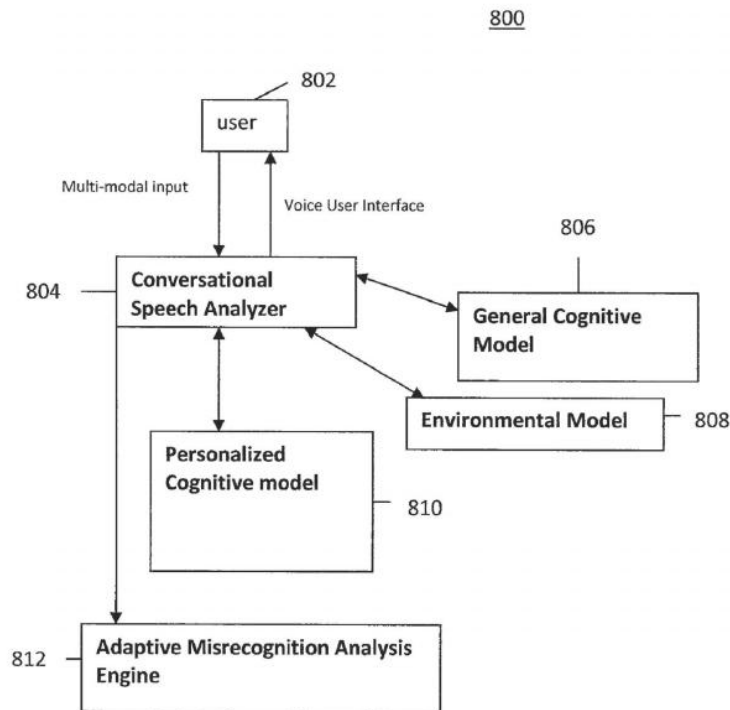


FIGURE 8

’468 patent, FIG. 8; *see also id.*, 15:45-47. In FIG. 8, an enhanced conversational user interface 800 receives multi-modal input from users 802 and communicates with a conversational speech analyzer 804. ’468 patent, 23:48-51. “The conversational speech analyzer 804 may be coupled to several components,

including a general cognitive model 806, an environmental model 808, a personalized cognitive model 810, and an adaptive misrecognition analysis engine 812 or other components.” ’468 patent, 23:51-55.

Data that is entered in a text format from a non-speech input may be merged with data that is transcribed to a textual format from the speech utterance. ’468 patent, 23:62-64. “[T]he conversational speech analyzer 804 may include one or more speech recognition engines that transcribe utterances to textual messages. The transcription may be performed using one or more conversational grammars, a context description grammar 112, or other transcription methods. ’468 patent, 23:56-61. “Conversational speech analyzer 804 also may include a semantic knowledge-based model that analyzes the textual message and detects command components.” ’468 patent, 23:64-67.

“[A] personalized cognitive model 810 is a model derived from a user’s interaction pattern with the system and may be used to predict what actions the user may take next in time, thus assisting with speech recognition and/or question or command recognition.” ’468 patent, 24:1-6.

“[A] general cognitive model 806 is a statistical abstract that corresponds to interaction patterns with the system for multiple users.” ’468 patent, 24:12-15.

Data stored within general cognitive model 806 may be used to “predict a user’s

next action,” thus assisting with speech recognition and/or question or command recognition. ’468 patent, 24:15-17.

“The conversational speech analyzer 804 may also access the general cognitive model 806 and/or the personalized cognitive model 810 to further refine a context, domain knowledge, preferences, and/or other cognitive qualities to enhance the interpretation of questions and/or commands.” ’468 patent, 24:40-44.

Claim 19 of the ’468 patent recites:

19. A method for processing multi-modal natural language inputs, comprising:

- receiving a multi-modal natural language input at a conversational voice user interface, the multi-modal input including a natural language utterance and a non-speech input provided by a user, wherein a transcription module coupled to the conversational voice user interface transcribes the non-speech input to create a non-speech-based transcription;

- identifying the user that provided the multi-modal input;

- creating a speech-based transcription of the natural language utterance using a speech recognition engine and a semantic knowledge-based model, wherein the semantic knowledge-based model includes a personalized cognitive model derived from one or more prior interactions between the identified user and the conversational voice user interface, a general cognitive model derived from one or more prior

interactions between a plurality of users and the conversational voice user interface, and an environmental model derived from an environment of the identified user and the conversational voice user interface; merging the speech-based transcription and the non-speech-based transcription to create a merged transcription;

identifying one or more entries in a context stack matching information contained in the merged transcription; determining a most likely context for the multi-modal input based on the identified entries;

identifying a domain agent associated with the most likely context for the multi-modal input; communicating a request to the identified domain agent; and generating a response to the user from content provided by the identified domain agent as a result of processing the request.

'468 patent, claim 19 (emphasis added).

III. THE BOARD NEED NOT REACH ANY CLAIM CONSTRUCTION ISSUES TO DENY THE PETITION

The Petitioner identifies no terms as requiring constructions. Without waiving its right to submit constructions in a Patent Owner Response, Patent Owner submits that the Board need not engage in any pre-institution claim construction. For the reasons presented in Section IV below, Petitioner's grounds include fundamental flaws and fail to establish the disclosure of every claimed element in each independent claim. These findings do not require the Board to conduct any construction of the above-identified terms.

IV. PETITIONER'S CHALLENGES FAIL ON THE MERITS

Petitioner asserts that independent claim 19 and dependent claims 25-30 and 32 are obvious in view of Maes² (Ground 1) and in view of Maes, Coffman,³ and Ittycheriah⁴ (Ground 2).

As introduced above, Petitioner's invalidity challenge specifically fails as to element [19.6]. Per element [19.5], the claimed method includes the step of "identifying one or more entries in a context stack matching information contained in the merged transcription," and element [19.6] includes the step of "determining a most likely context [in the context stack] for the multimodal input based on the identified entries." For these elements, Petitioner relies on Maes's teaching of "scoring one or more candidate contexts such as phonemes," which are described by Maes as "sub-phonetic or acoustic units of speech." Pet., 62-63; Maes, 12:31-33. But Petitioners fail to realize that Maes's scoring candidate contexts is about trying to determine what a user has said, and not whether the user's utterance matches a context in a context stack. *See* Maes, 12:35-37 ("Each phoneme associated with one or more feature vectors has a probability associated therewith ***indicating the likelihood that it was that particular acoustic unit that was***

² Ex. 1005, U.S. Patent No. 6,964,023 ("Maes").

³ Ex. 1012, International Patent Application Publication No. WO 00/20962 ("Coffman").

⁴ Ex. 1006, U.S. Patent No. 5,937,383 ("Ittycheriah").

spoken.”) (emphasis added). Maes confirms this purpose of phonemes, which is discussed in 12:29-48 in relation to Fig. 4, a block diagram of an “audio-visual speech recognition module” (3:33-34) whose purpose is to recognize what the user has said. Thus, Maes’s process of trying to determine whether a user’s utterance matches “phonemes” is about trying to determine what the user uttered. It is not relevant to the claimed steps of “identifying one or more entries in a context stack matching information contained in the merged transcription,” (element [19.5]) and “determining a most likely context for the multimodal input based on the identified entries” (element [19.6]). The Petition is completely wrong in its analysis, particularly with regard to element [19.6], because determining what a user has uttered is not the same as determining a mostly likely context in a context stack to match that utterance.

Claim 19, and through dependency each of the challenged claims, also recites a “semantic knowledge-based model” that includes “a personalized cognitive model” and “a general cognitive model.” In addition to the Petition’s shortcoming as to element [19.6], Petitioner also fails to establish that the asserted references include such “cognitive” models.

Instead, Petitioner follows the lead of its expert and submits an altered quotation of the ’468 patent to suggest that a cognitive model merely “assists” with speech recognition. Pet., 36. But that is not what the ’468 patent says and is not

what the claims require. Specifically, Petitioner's expert begins his analysis of the "cognitive" elements with the statement, Ex. 1003, ¶146 (highlighting added):

146. Initially, the 468 Patent describes that a "*cognitive model*" as "a model derived from a user's interaction pattern with the system and may be used to predict what actions the user may take next in time [and] assisting with speech recognition and/or question or command recognition." Ex.1001, 24:1-11.

But the expert's alteration of the quotation changes the meaning of how a POSITA would understand a "cognitive" model in the context of the '468 patent. Further, as discussed below, in analyzing this aspect of the challenged claims, Petitioner's expert focused solely on whether a model would assist with speech recognition, without addressing whether such model was properly characterized as "cognitive."

Accurately quoted, the '468 patent says with regard to a "personalized cognitive model" ('468 patent, 24:1-11) (emphasis added):

may be used to predict what actions the user may take
next in time, *thus* assisting with speech recognition
and/or question or command recognition

Similarly, for a "general cognitive model," the '468 patent states that data in the model "may be used to predict a user's next action, thus assisting with speech recognition and/or question or command recognition." '468 patent, 24:12-17.

The expert's seemingly subtle alteration of the '468 patent by substituting "and" for " , thus" changes the meaning of the passage, and the expert relied upon this

slight change to justify his characterization of any model used to “assist” speech recognition to satisfy the claimed “cognitive” model requirement. *See Merck & Co. v. Teva Pharm. 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.” (citations omitted)). This was error.

Specifically, Petitioner’s expert looked for “personal” models without opining whether they would meet the claimed “cognitive” requirement. Paragraph 147 of his analysis speaks to models that “assist in speech recognition.” Ex. 1003, ¶147. Paragraph 148 looks to a reference that is not included in Petitioner’s Grounds and does not address the “cognitive” requirement. Ex. 1003, ¶148. Paragraphs 149-155 look to the Maes reference and, other than conclusory statements, fail to provide another Petitioner’s understanding of “cognitive” or how Maes satisfies the “cognitive” requirement. Ex. 1003, ¶¶149-55.

Perhaps recognizing the deficiencies of Maes, the expert then turned to Ittycheriah, purportedly to address the “cognitive” model aspect. Ex. 1003, ¶156. But rather than explain how Ittycheriah satisfies the “cognitive” aspects, the expert addressed how a cache database in Ittycheriah contains user-specific data and how the data is “used to test a hypothesis.” Ex. 1003, ¶¶157-58. Nowhere in that discussion does the expert explain what makes Ittycheriah’s cache database a

“cognitive” model or even opine about such a characterization. For example, there certainly is no discussion of how Ittycheriah’s cache database “predicts a user’s next action.”

The expert’s analysis of the “general cognitive model” is no different and similarly fails to show how the references provide a “cognitive” model. Instead, the expert incorporates his prior defective analysis and addresses “multi-speaker” aspects of Maes, all the while failing to tackle the “cognitive” requirement. Ex. 1003, ¶¶161-68.

The Petition fares no better, and merely paraphrases the declaration of Petitioner’s expert without correcting the mistaken quotation (Pet., 36) or deficiencies of its expert’s analysis. *See generally* Pet. 36-43. Accordingly, Petitioner and its expert fail to show how the claimed aspects of a “cognitive model” should be construed or how the references applied in its Grounds satisfy those aspects. These deficiencies render the challenges insufficient at least as to the claimed “cognitive model” requirements, and warrant denial of the Petition.

Finally, Ground 2 adds Coffman and Ittycheriah to the Maes challenge of Ground 1 in case “one asserts that Coffman and Ittycheriah are not particularly incorporated by reference into Maes as if they were effectively the same document.” Pet., 80. However, and as noted above, element [19.6] includes the step of “determining a most likely context [in the context stack] for the multimodal

input based on the identified entries,” and Petitioner relies solely on Maes for this element. Pet., 62-63. Coffman's and Ittycheriah's other disclosures are not relied upon to remedy the above-discussed deficiencies of Maes as to this element. As such, the failures of Ground 1 are imputed to Ground 2 as well, and the Petition establishes no basis for institution.

V. CONCLUSION

For the reasons presented above, the Petition should be denied, and no *inter partes* review should be instituted.

Dated: July 15, 2025

/Wayne M. Helge/

Wayne M. Helge, Reg. No. 56,905

James T. Wilson, Reg. No. 41,439

BUNSOW DE MORY LLP

277 S. Washington St., Suite 210 #1088

Alexandria, VA 22314

T: (571) 208-0186

Email: whelge@bdiplaw.com

Email : jwilson@bdiplaw.com

Counsel for Patent Owner

CERTIFICATE OF WORD COUNT

The undersigned certifies that the foregoing PATENT OWNER'S PRELIMINARY RESPONSE complies with the type-volume limitation in 37 C.F.R. § 42.24(b)(1). According to the word-processing system's word count, the brief contains 2,807 words, excluding the parts of the brief exempted by 37 C.F.R. § 42.24(a).

By: /Wayne M. Helge/
Wayne M. Helge (Reg. No. 56,905)
Counsel for Patent Owner

CERTIFICATE OF SERVICE

I hereby certify that on July 15, 2025, a true and correct copy of the foregoing document was served via email, by consent, to Petitioner by serving the correspondence email addresses of record as follows:

WINSTON & STRAWN LLP
Alias (Winston-Microsoft-Dialect-IPR@winston.com)
Scott M. Border (sborder@winston.com)
Carson Swope (cswope@winston.com)
T.D. Goswami (tdgoswami@winston.com)

/Wayne M. Helge/
Wayne M. Helge (Reg. No. 56,905)
Counsel for Patent Owner