UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.,
Petitioner,

v.

DIALECT, LLC,
Patent Owner.

Case IPR2025-01332
Patent 8,447,607

PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I. INTRODUCTION 1

II. THE '607 PATENT 2

III. THE BOARD NEED NOT REACH ANY CLAIM CONSTRUCTION ISSUES TO DENY THE PETITION 9

IV. PETITIONER'S CHALLENGES FAIL ON THE MERITS 9

V. CONCLUSION 14

PATENT OWNER'S EXHIBIT LIST

Exhibit No.	Description
2001	Decision Denying Institution of *Inter Partes* Review, *Microsoft Corporation v. Dialect, LLC*, IPR2025-00656, Paper 12 (PTAB Aug. 14, 2025)
2002	Patent Owner Preliminary Response, *Microsoft Corporation v. Dialect, LLC*, IPR2025-00656, Paper 7 (PTAB July 15, 2025)
2003	Defendant's Preliminary Invalidity Contentions, *Dialect, LLC v. Meta Platforms, Inc.*, Civ. No. 7:25-cv-60 (W.D. Tex. Sept. 17, 2025)
2004	Redacted Order in *Dialect, LLC v. Samsung Electronics Co., Ltd.*, Civ. No. 2-23-cv-00061 (E.D. Tex. Aug. 28, 2023)
2005	Order Acknowledging Stipulation of Dismissal With Prejudice, *Dialect, LLC v. Amazon.com, Inc.*, Civ. No. 1:23-cv-00581 (E.D. Va. Jan. 23, 2025)
2006	Second Amended Docket Control Order in *Dialect, LLC v. Bank of America, N.A.*, Civ. No. 2:24-cv-00207 (E.D. Tex. June 27, 2025)
2007	Decision Denying Institution of *Inter Partes* Review, *Google LLC v. Dialect, LLC*, IPR2024-00751, Paper 11 (PTAB Oct. 11, 2024)

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 12 and 13 of U.S. Patent No. 8,447,607 (the "'607 patent") , as requested by Meta Platforms, Inc. ("Petitioner" or "Meta"), should be denied.[1]

I. INTRODUCTION

The '607 patent claims a device that generates speech-based transcriptions using a cognitive model associated with an identified user, where that cognitive model includes information from prior interactions between the user and the device. Petitioner's challenges to the '607 patent should be denied because the asserted prior art references fail to disclose these aspects of the claimed invention.

In place of teachings in the asserted prior art, the Petition turns to an improper reliance on speculation from its expert to manufacture the missing claim elements. Specifically, rather than pointing to specific disclosures in the Maes, Coffman, and Ittycheriah references, Petitioner instead relies on its expert to theorize about general language model training processes that bear no relationship to the claimed user-specific cognitive models based on prior interactions. Even where the references discuss dynamic user attribute determination, these features

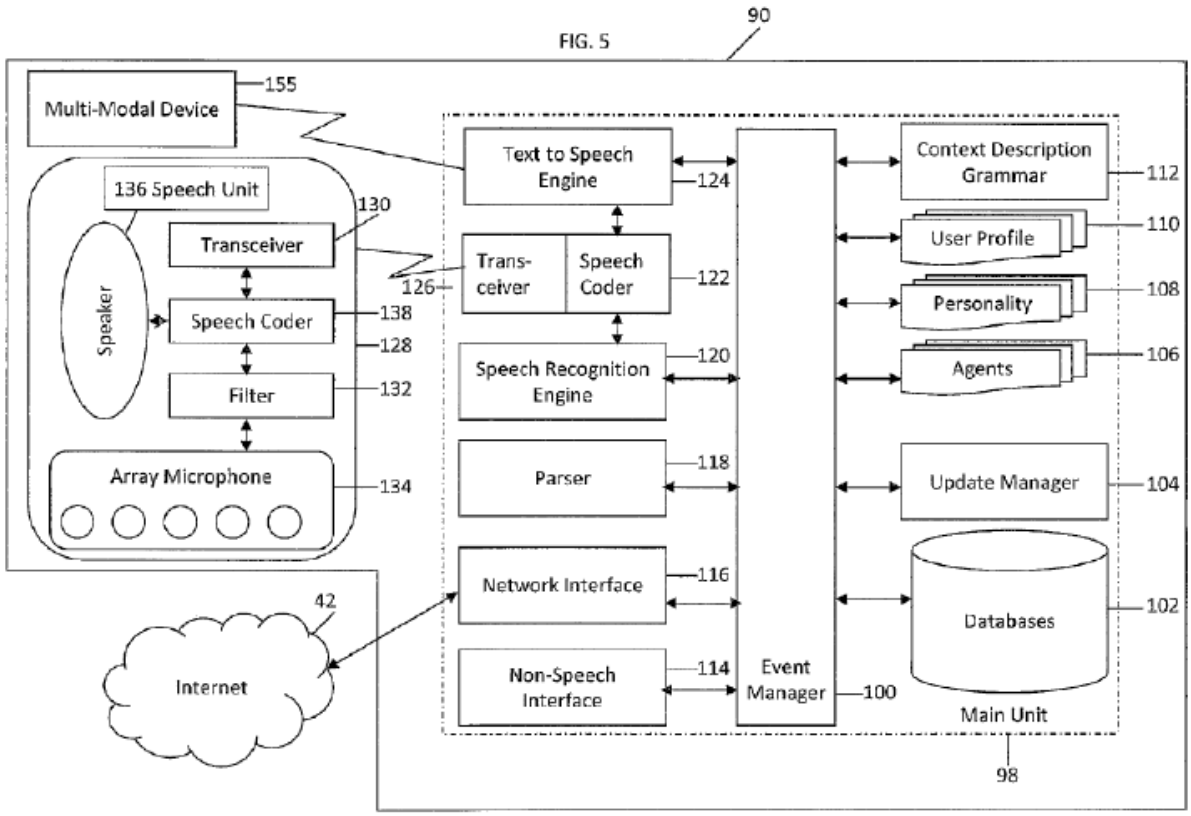
[1] By submitting this Preliminary Response, no waiver of any argument is intended by Patent Owner. Patent Owner will have a right to file a complete response to the Petition, should the Board institute *inter partes* review. 37 C.F.R. § 42.120(a).

operate during current interactions—***not based on stored information from previous user sessions as required by the claims***. Petitioner's repeated attempts to characterize databases, language models, and speaker identification systems as the claimed "cognitive model" reflect a fundamental misunderstanding of the claimed invention and constitute impermissible expert-driven gap-filling. Each of Petitioner's obviousness-based theories fails because the prior art simply does not teach the claimed cognitive model functionality.

II. THE '607 PATENT

The '607 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). Central to the invention claimed in the '607 patent, the '607 patent enables merging of natural language speech and non-speech questions or commands to identify the proper context for the input and use of a cognitive model associated with the user.

FIG. 5 of the '607 patent, reproduced below, is an overall diagrammatic view of an embodiment of the '607 patent:



'607 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. Speech unit 128 includes a microphone to receive a spoken utterance from a user. '607 patent, 25:50-52. After initial processing, the speech is passed to speech recognition engine 120 for processing using the context description grammar module 112. '607 patent, 26:18-20. 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." '607 patent, 26:20-23. "The knowledge agents

may then process the commands or questions." '607 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." '607 patent, 26:51-53.

FIG. 8 of the '607 patent, reproduced below, illustrates aspects of an embodiment of the '607 patent in more detail.

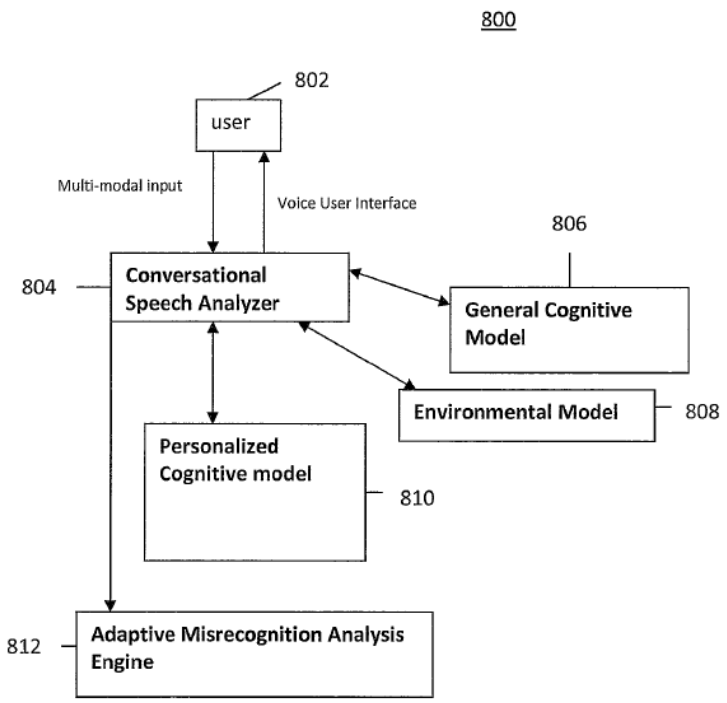


FIGURE 8

In FIG. 8, an enhanced conversational user interface 800 "receives multi-modal input from users 802 and communicates with a conversational speech analyzer 804." '607 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." '607 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. '607 patent, 23:62-64. The 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. '607 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." '607 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." '607 patent, 24:1-6.

"[A] general cognitive model 806 is a statistical abstract that corresponds to interaction patterns with the system for multiple users. '607 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. '607 patent, 24:15-17.

An environmental model 808 may include information associated with the user's environment and surroundings. The information may include the type of environment that a user is in (e.g., quiet or noisy); details of a microphone and/or speaker system; the user's current global position and movement, as may be determined by GPS; current system status, such as what song/movie is playing, is the system in the midst of retrieving something, or other system status; details on all voice-enabled devices in the immediate vicinity, such as a presence of a voice-enabled TV, stereo, and DVD player in the same room; a user's credit card information, such as numbers and current balances, wherein the user may ask a mobile telephone to download and pay for a video and the system may respond that there are insufficient funds; or other information. '607 patent, 24:22-36. "The information may be accessed to invoke a context, domain knowledge, preferences, and/or other cognitive qualities that enhance the interpretation of questions and/or commands." '607 patent, 24:37-39.

"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." '607 patent, 24:40-44.

As described in the '607 patent, context information may be determined from a command or request that is presented in a textual format and/or a command

or request that is presented as an utterance and processed using a multi-pass automatic speech recognition module that transcribes the utterance to a text message. '607 patent, 3:66-4:2.

Consistent with the '607 patent's disclosure, independent claims 1 and 12 of the '607 patent require a multi-modal natural language input produced by merging a speech transcription and non-speech transcription and use of that merged transcription to identify an entry in a context stack. Each independent claim also recites use of a cognitive model associated with the user.

Claim 12 recites:

> 12. A method for processing natural language inputs, comprising:
>
> receiving, by one or more processors, a multi-modal natural language input from a user, the multi-modal natural language input including a natural language utterance and a non-speech input;
>
> generating, by the one or more processors, a non-speech transcription from the non-speech input;
>
> ***identifying***, by the one or more processors, ***the user who provided the multi-modal natural language input***;
>
> ***generating***, by the one or more processors, ***a speech-based transcription based on a cognitive model associated with the user, wherein the cognitive model includes information on one or more prior interactions between the user and the device***;

generating, by the one or more processors, a merged transcription from the speech-based transcription and the non-speech transcription;

identifying, by the one or more processors, an entry in a context stack, from among a plurality of entries that are in the context stack and that are each indicative of context, wherein the identified entry matches information in the merged transcription;

identifying, by the one or more processors, a domain agent associated with the entry in the context stack;

determining, by the one or more processors, a request based on the merged transcription; and

communicating, by the one or more processors, the request to the domain agent, wherein the domain agent is configured to generate a response to the user.

'607 patent, claim 12 (emphasis added).

Dependent claim 13 recites:

13. The method of claim 12, wherein the cognitive model is a first cognitive model, and wherein the generating the speech-based transcription is based on a second cognitive model that includes information on one or more interactions between the device and a plurality of users of the device.

'607 patent, claim 13.

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

The Petitioner identifies no terms as requiring constructions. Pet., 6. 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 assert that claims 12 and 13 are obvious in view of Maes[2] (Ground 1) and in view of Maes, Coffman,[3] and Ittycheriah[4] (Ground 2).

As noted above, the '607 patent relates to enabling natural language human-machine interactions with multimodal inputs, such as from a speech interface and nonspeech interfaces. As part of the claimed device, independent claim 12 of the '607 Patent requires the device to, *inter alia*, "identify the user who provided the multi-modal natural language input" [element 12.3] and "generate a speech-based transcription based on a cognitive model associated with the user, wherein the

[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").

cognitive model includes information on one or more prior interactions between the user and the device" [element 12.4]. Thus, the user must be identified, and the speech-based transcription is based on a cognitive model associated with that user. Further, the cognitive model requires information on at least one prior interaction between the user and the device. These elements are plainly missing from the prior art.

The Petition relies on the asserted references Maes, Coffman, and Ittycheriah to disclose or render obvious element [12.4] (*see* Pet., 35-48), but these references fail to disclose or suggest that any "cognitive model" allegedly disclosed "includes information on one or more prior interactions between the user and the device." Rather than pointing to these features in the asserted references, Petitioner instead relies on its expert to manufacture these features. Pet., 43 (relying on Dr. Houh's declaration to explain "that a POSITA would have understood that in the context of speech recognition systems, training a language model involves prior interactions between a user and a system… .").[5] Of course, even this sentence fails to tie "prior interactions between a user and a system" to

[5] The Petition briefly discusses Ex. 1008 to "Lee," but does not expressly include Lee in its challenge, for good reason. Pet., 41. Lee includes a database of previously-recognized words but it is separate from any "language models" and would not aid Petitioner's challenge. Lee, [0016]; FIG. 1 (indicating RECOGNITION ASSISTING DATABASE(S) 48 as a separate module and separate flow from LANGUAGE MODEL 36).

the user identified as providing the multi-modal natural language input, as required by element [12.3].

Indeed, even if Dr. Houh is to be believed on the process of training a language model, none of the asserted references disclose or suggest the process of generating "a speech-based transcription based on a cognitive model associated with the user," where the "cognitive model includes information on one or more prior interactions between the user and the device." While Petitioner asserts that Maes includes a "dynamic classification module" that can dynamically determine an acoustic feature such as accent, dialect, or native language of the user, (*see* Pet., 42), Maes does not disclose that any speech-based transcription is generated based on a prior interaction between a user and the device or even a prior determined acoustic feature such as accent, dialect, or native language of the user. Rather, those features are dynamically determined and extracted while the user is currently speaking, i.e., based on a user's ***current interaction***. *See* Maes, 44:7-11 ("processing module 1010 is configured to modify behavior of the voice system 1000 based on at least one user attribute which has been determined based on at least one acoustic feature extracted from the user's speech."); 44:22-25 ("Dynamic classification module 1040 can be configured to provide feedback which can be ***real-time feedback*** to the business logic unit 1042 and the dialog management unit 1002.") (emphasis added).

Petitioner next relies on its expert again to reinvent Maes by asserting that Maes's "teachings of employing a speaker for creating a model of a native language of that user" satisfies these claimed features. Pet., 43 (citing Dr. Houh, ¶122). But the mere creation of a model of a user's language does not mean that Maes is somehow using that same user's model to generate a speech-based transcription based on a cognitive model that includes information on one or more prior interactions between the user and the device. The Petition plainly fails to identify such features related to this native language model. Pet., 42-43. Surely if Maes disclosed such features, Petitioner would not have needed to spend nearly 13 pages trying to build a case of obviousness as to this feature. *See* Pet., 35-48.

Petitioner's remaining arguments of an alleged cognitive model in Maes that includes prior interactions between the device and the identified user who provided the "multi-modal natural language input" [per element 12.3] are equally lacking. The Petition points to "language model training" in Maes (Pet., 43), but fails to connect such training to one or more prior interactions with the very user who has been identified as providing the "multi-modal natural language input" in question. Even the disclosed "speaker profile features (accent, dialect, gender, age, speaking rate, identity, audio quality, SNR, etc . . .)" (Maes, 19:18-20) are not identified as characteristics of a ***current speaker based on a prior interaction***.

Moreover, Petitioner's last-ditch effort of Maes's "stored models" (Pet., 44) is also off-point. Maes discloses that these stored "models can be selected to be similar to the claimant speaker or be speaker independent" (Maes, 24:31-45) but Maes discloses that the speaker model is used to ***determine whose utterance is being processed*** so that the speaker's audio and video data can be compared, and not to generate a speech-based transcription as required by element [12.4]. Pet., 44-45 (quoting Maes, 24:28-30) ("the above approach 'finds the closest matching model and the person whose model that represents is determined to be the person whose utterance is being processed.'") (emphasis omitted). The goal of speaker identification, in Maes, is not to retrieve a prior utterance from that same speaker, but rather to "to find the model, M_i, that best explains the test data represented by a sequence of N frames … ." Maes, 24:11-14.

Petitioner reverts to Ittycheriah (U.S. Pat. No. 5,937,383), but this reference also falls short. Pet., 45-47. As disclosed by Ittycheriah, the system partitions a database to store individual speakers' inputs separately, but establishing a separate database falls short of a "cognitive model" as required by element [12.4]. At best, Petitioner uses sleight-of-hand to refer to "such database used as a model" (Pet., 45) and characterizes a "cache database" as "a cognitive model" as claimed. Pet., 46-47. These characterizations are unsupported and fall short. Once again, the Petition fails to carry its burden as to this claimed feature.

Finally, Ground 2 adds Coffman and Ittycheriah to the Maes challenge of Ground 1 in case "one asserts that Coffman and Ittycheriah are not expressly and/or particularly incorporated by reference into Maes as if they were effectively the same document." Pet., 78. However, and as noted above, Petitioner relies on Maes and Ittycheriah for element [12.4]. Pet., 35-48. Coffman's disclosures are not relied upon to remedy the above-discussed deficiencies of Maes and Ittycheriah as to this element [12.4]. 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: October 31, 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 2,753 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 October 31, 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:

PAUL HASTINGS LLP
Lisa K. Nguyen
David Tennant
Diane Ghrist
Rachel Wu Hankinson
PH-META-DialectIPR@paulhastings.com

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