

Petition for *Inter Partes* Review of
U.S. Patent No. 9,031,845 B2

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

GOOGLE LLC,
Petitioner

v.

DIALECT, LLC,
Patent Owner

Case IPR2024-00750
U.S. Patent No. 9,031,845 B2
Issue Date: May 12, 2015

Title: MOBILE SYSTEMS AND METHODS FOR RESPONDING TO
NATURAL LANGUAGE SPEECH UTTERANCE

**PETITION FOR *INTER PARTES* REVIEW
OF U.S. PATENT NO. 9,031,845 B2**

Table of Contents

	Page
I. MANDATORY NOTICES UNDER §42.8(A)(1)	8
A. Real Party-In-Interest under §42.8(b)(1).....	8
B. Related Matters under §42.8(b)(2)	8
C. Lead and Back-Up Counsel under §42.8(b)(3)	9
D. Service Information	10
II. FEE PAYMENT	10
III. REQUIREMENTS UNDER §§ 42.104 AND 42.108 AND CONSIDERATIONS UNDER §§ 314(A) AND 325(D).....	10
A. Grounds for Standing	10
B. Identification of Challenge and Statement of Precise Relief Requested	11
C. Considerations Under §§ 314(a) and 325(d).....	11
IV. OVERVIEW OF THE CHALLENGED PATENT.....	15
A. Person of Ordinary Skill in the Art (POSITA)	15
B. Specification Overview	16
V. CLAIM CONSTRUCTION	17
VI. THE CHALLENGED CLAIMS ARE UNPATENTABLE.....	17
A. Overview of Grounds	17
B. Brief Summary of Prior Art	18
1. Coffman (EX1003)	18
2. Julia (EX1004)	21
3. Cooper (EX1005)	23
C. Ground 1: Claims 1-3, 7-8, 10, 12, 15, 17-20, 24-25, 27, 29, 32, 34, 36 Are Obvious Over Coffman In View Of Julia And Cooper.	23
1. Claim 1: “A mobile system for processing natural language utterances, comprising:” (1[pre])	23

Table of Contents

(continued)

	Page
(a) “one or more physical processors at a vehicle that are programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:”	24
(b) “receive a natural language utterance associated with a user;”	25
(c) “perform speech recognition on the natural language utterance;”	37
(d) “parse and interpret the speech recognized natural language utterance;”	37
(e) “determine a domain and a context that are associated with the parsed and interpreted natural language utterance;”	39
(f) “formulate a command or query based on the domain and the context;”	43
(g) “determine whether the command or query is to be executed on-board or off-board the vehicle;”	45
(h) “execute the command or query at the vehicle in response to a determination that the command or query is to be executed on-board the vehicle; and”	51
(i) “invoke a device that communicates wirelessly over a wide area network to process the command or query such that the command or query is executed off-board the vehicle in response to a determination that the command or query is to be executed off-board the vehicle.”	52
2. Claim 2: “The mobile system of claim 1, wherein executing the command or query at the vehicle comprises executing the command or query at a device built into the vehicle.”	55

Table of Contents
(continued)

Page

3.	Claim 3: “The mobile system of claim 1, wherein executing the command or query at the vehicle comprises executing the command or query at a device docked to the vehicle.”	56
4.	Claim 7: “The mobile system of claim 1, wherein executing the command or query at the vehicle comprises executing the command or query at a device that is in or on the vehicle and that is in wireless communication with the vehicle.”	57
5.	Claim 8: “The mobile system of claim 1, wherein the command or query includes a search query that is to be executed on-board the vehicle by searching one or more data sources that are on-board the vehicle using at least one recognized word or phrase of the natural language utterance, and wherein the search query relates to an aspect of the vehicle.”	58
6.	Claim 10: “The mobile system of claim 8, wherein the instructions cause the one or more physical processors to:” (10[pre])	61
	(a) “provide a natural language speech response to the natural language utterance based on one or more results from the search of the one or more data sources.” (10[a])	62
7.	Claim 12: “The mobile system of claim 1, wherein the command or query is to be executed on-board the vehicle, and wherein the command or query includes a command to control the vehicle.”	64
8.	Claim 15: “The mobile system of claim 1, wherein the instructions cause the one or more physical processors to:” (15[pre])	65
	(a) “identify a context associated with the natural language utterance based on a current or prior dialog between the system and the user,” (15[a])	65

Table of Contents

(continued)

	Page
(b) “wherein the determination of whether the command or query is to be executed on-board or off-board the vehicle is further based on the context.” (15[b])	66
9. Claim 17: “The mobile system of claim 1, wherein the command or query includes a command that is to be executed at the vehicle, and wherein the instructions cause the one or more physical processors to:” (17[pre]).....	67
(a) “identify a query based on the natural language utterance;” (17[a]).....	68
(b) “invoke a device that communicates wirelessly over a wide area network to process the query such that the query is executed off-board the vehicle based on a determination that the query is to be executed off-board the vehicle; and” (17[b])	68
(c) “receive, from a device off-board the vehicle, a response to the query, wherein the command is executed at the vehicle based on the query response.” (17[c])	69
10. Independent Claim 18	70
11. Claims 19, 20, 24, 25, 27 & 29	70
12. Claim 32: “A computer-implemented method of processing natural language utterances, the method being implemented by a computer system that includes one or more physical processors at a vehicle executing one or more computer program instructions which, when executed, perform the method, the method comprising:” (32[pre])	70
(a) “receiving, at the one or more physical processors, a natural language utterance associated with a user;” (32[a]).....	71

Table of Contents (continued)

Page

(b)	“performing, by the one or more physical processors, speech recognition on the natural language utterance;” (32[b]).....	71
(c)	“parsing and interpreting, by the one or more physical processors, the speech recognized natural language utterance;” (32[c])	71
(d)	“determining, by the one or more physical processors, a domain and a context that are associated with the parsed and interpreted natural language utterance;” (32[d]).....	71
(e)	“formulating, by the one or more physical processors, a search query based on the domain and the context wherein the search query includes at least one recognized word or phrase of the natural language utterance;” (32[e]).....	72
(f)	“determining, by the one or more physical processors, whether the search query is to be executed on-board or off-board a vehicle;” (32[f])	76
(g)	“executing, by the one or more physical processors, the search query at the vehicle in response to a determination that the search query is to be executed on-board the vehicle; and” (32[g])	77
(h)	“invoking, by the one or more physical processors, a device that communicates wirelessly over a wide area network to process the search query such that the search query is executed off-board the vehicle in response to a determination that the search query is to be executed off-board the vehicle.” (32[h])	78
13.	Claims 34 & 36	79
VII.	CONCLUSION.....	79
	CERTIFICATE OF SERVICE	81

Petition for *Inter Partes* Review of
U.S. Patent No. 9,031,845 B2

List of Exhibits

Exhibit No.	Description of Document
1001	U.S. Patent No. 9,031,845 B2 to Robert A. Kennewick et al. (filed February 12, 2010, issued May 12, 2015) (“ 845 ” or “ 845 patent ”)
1002	Declaration of Christopher M. Schmandt (“ Schmandt ”)
1003	International Patent Application Publication No. WO 00/20962 to Daniel Coffman, Stephane H. Maes et al. (filed October 1, 1999, published April 13, 2000) (“ Coffman ”)
1004	U.S. Patent No. 7,036,128 B1 to Luc Julia et al. (filed August 9, 2000, issued April 25, 2006) (“ Julia ”)
1005	U.S. Patent No. 6,757,362 B1 to Robert S. Cooper et al. (filed March 6, 2000, issued June 29, 2004) (“ Cooper ”)
1006	U.S. Patent No. 6,801,994 B2 to Richard Dennis Beckert et al. (filed December 20, 2000, issued October 5, 2004) (“ Beckert ”)
1007	<i>RESERVED</i>
1008	Excerpts from Dafydd Gibbon et al., <i>Handbook of Multimodal and Spoken Dialogue Systems : Resources, Terminology, and Product Evaluation</i> (2000) (“ Gibbon ”)
1009	Excerpts from <i>Microsoft Computer Dictionary</i> (5th ed. 2002)
1010-1017	<i>RESERVED</i>
1018	Declaration of Ingrid Hsieh-Yee, Ph.D.
1019	Prosecution history of U.S. Patent No. 9,031,845 B2
1020	Proof of Service of Complaint, <i>Dialect, LLC v. Google LLC</i> , Case No. 1:23-cv-00378 (D. Del.)

I. MANDATORY NOTICES UNDER §42.8(A)(1)

A. Real Party-In-Interest under §42.8(b)(1)

Google LLC (“Petitioner”) is the real party-in-interest to this IPR petition.¹

B. Related Matters under §42.8(b)(2)

The ’845 patent is the subject of pending litigation in which Petitioner is named as a defendant, *Dialect, LLC v. Google LLC*, Case No. 1:23-cv-00378 (D. Del.) (filed April 3, 2023). Petitioner was first served with a complaint alleging infringement of the ’845 patent on April 6, 2023. (EX1020, p.002.) On March 28, 2024, Petitioner’s motion to transfer to the Northern District of California was granted. *See* Dkt. 32, Case No. 1:23-cv-00378 (D. Del.). The ’845 patent is also the subject of *Dialect, LLC v. Amazon.com, Inc.*, Case No. 1:23-cv-00581-LMB-LRV (E.D. Va.) (filed May 1, 2023). In the *Amazon.com* litigation, claim 1 of the ’845 patent was held invalid under 35 U.S.C. § 101. *See Dialect v. Amazon.com*, Dkt. 58 (Nov. 8, 2023).

Petitioner is filing IPR petitions against eight patents asserted by Patent Owner in the pending litigation involving Petitioner:

¹ Google LLC is a subsidiary of XXVI Holdings Inc., which is a subsidiary of Alphabet Inc. XXVI Holdings Inc. and Alphabet Inc. are not real parties in interest to this proceeding.

Petition for *Inter Partes* Review of
U.S. Patent No. 9,031,845 B2

U.S. Patent No.	IPR Case
7,398,209 B2	IPR2024-00746
7,502,738 B2	IPR2024-00747
8,015,006 B2	IPR2024-00748
7,693,720 B2	IPR2024-00749
9,031,845 B2	IPR2024-00750
8,447,607 B2	IPR2024-00751
8,849,652 B2	IPR2024-00752
7,640,160 B2	IPR2024-00753

C. Lead and Back-Up Counsel under §42.8(b)(3)

Petitioner provides the following designation of counsel.

LEAD COUNSEL	BACK-UP COUNSEL
Heidi L. Keefe (Reg. No. 40,673) hkeefe@cooley.com COOLEY LLP ATTN: Patent Group 1299 Pennsylvania Ave. NW, Suite 700 Washington, DC 20004 Tel: (650) 843-5001 Fax: (650) 849-7400	Andrew C. Mace (Reg. No. 63,342) amace@cooley.com COOLEY LLP ATTN: Patent Group 1299 Pennsylvania Ave. NW, Suite 700 Washington D.C. 20004 Tel: (650) 843-5287 Fax: (650) 849-7400
	Mark R. Weinstein (Admission <i>pro hac vice</i> to be requested) mweinstein@cooley.com COOLEY LLP ATTN: Patent Group 1299 Pennsylvania Avenue NW Suite 700 Washington D.C. 20004 Tel: (650) 843-5007 Fax: (650) 849-7400

LEAD COUNSEL	BACK-UP COUNSEL
	Reuben Chen (Admission <i>pro hac vice</i> to be requested) rchen@cooley.com COOLEY LLP ATTN: Patent Group 1299 Pennsylvania Avenue NW Suite 700 Washington D.C. 20004 Tel: (650) 843-5007 Fax: (650) 849-7400

D. Service Information

This Petition is being served by Federal Express to the correspondence address of record for the '845 patent, 174899 - The Law Offices of David A. Gerasimow, P.C., 211 W. Wacker Dr., Ste. 1717, Chicago, IL 60606. Petitioner consents to electronic service at the addresses provided above for lead and back-up counsel.

II. FEE PAYMENT

Petitioner requests review of 19 claims, with a \$41,500 payment.

**III. REQUIREMENTS UNDER §§ 42.104 AND 42.108 AND
CONSIDERATIONS UNDER §§ 314(A) AND 325(D)**

A. Grounds for Standing

Petitioner certifies that the '845 patent is available for IPR and that Petitioner is not barred or otherwise estopped.

B. Identification of Challenge and Statement of Precise Relief Requested

Petitioner requests IPR institution based on:

Ground	Claims	Basis for Challenge under §103
1	1-3, 7-8, 10, 12, 15, 17-20, 24-25, 27, 29, 32, 34, 36	Coffman (EX1003) in view of Julia (EX1004) and Cooper (EX1005)

The references are prior art under at least pre-AIA §§102(b) (Coffman) and 102(e) (Julia, Cooper).

C. Considerations Under §§ 314(a) and 325(d)

§314(a): Petitioner is unaware of another post-grant challenge against the '845 patent. The pending litigation against Petitioner is still in early stages, with no discovery or claim construction having occurred and no trial date set, and as noted, was recently transferred to the Northern District of California. If IPR is instituted, Petitioner intends to bring a motion to stay the litigation.

§325(d): During prosecution, pending claims were initially repeatedly rejected as obvious, and Coffman's U.S. patent counterpart (U.S. 7,137,126) and a related Coffman U.S. patent application publication (U.S. 2009/0313026) were identified as "pertinent to the applicant's disclosure." (EX1019, pp.00809-00834 (rejection), 00833-00834 (Coffman), 00854-00871 (claims), 00945-00983 (rejection), 00981-00983 (Coffman), 001062-001086 (claims), 001128-001198

(rejection), 001194-001198 (Coffman).) After those rounds of amended and new claims followed by Examiner rejections, the applicant filed amended and new claims that were rejected as anticipated and obvious over Coffman's U.S. counterpart. (EX1019, pp.001252-001270 (claims), 001275-001309 (rejection).) Among other things, the Examiner found that Coffman disclosed a claimed step of determining whether a command or query related to a natural language utterance is to be processed on-board or off-board the vehicle. (EX1019, pp.001281.) Following an applicant-initiated interview (EX1019, p.001360), the applicant further amended the pending claims (EX1019, pp.001363-001374), with the independent claims primarily amended to recite that the step of determining whether a command or query related to a natural language utterance is to be processed on-board or off-board the vehicle is performed "at a device that is on-board the vehicle" (*see* EX1019, pp.001364, 001368, 001372). The applicant stated that "[a]s agreed during the Interview, the references cited in the Final Action do not teach or suggested independent claims 56, 71, and 118 as currently amended." (EX1019, p.001373 (emphasis in original).) In response, the Examiner in an advisory action stated that U.S. Patent No. 6,721,633 to Funk taught the added limitations. (EX1019, pp.001380-001382.) The Examiner did not indicate why rejections based on Coffman's U.S. counterpart were not being maintained or otherwise discuss the reference. The applicant then further amended the pending claims to attempt

overcome the Funk reference. (EX1019, pp.001439-001459.) A notice of allowability followed, where the pending claims were allowed subject to an Examiner's Amendment. (EX1019, pp.001702-001713.) Similar to the "determining" step performed "at a device that is on-board the vehicle" noted above, the allowed independent claims recited a step of "determin[ing] [by one or more physical processors at a vehicle] whether the command or query is to be executed on-board or off-board the vehicle." (EX1019, pp.001704-001706, 001708-001709.)

Petitioner respectfully submits that the Examiner materially erred by failing to reject the allowed claims as obvious over Coffman. As noted, the Examiner found that Coffman disclosed a claimed step of determining whether a command or query related to a natural language utterance is to be processed on-board or off-board the vehicle. (EX1019, pp.001281.) But apparently, according to the applicant, the Examiner agreed that Coffman did not teach or suggest performing such a step *at a vehicle*. (EX1019, pp.001373, 001704-001706, 001708-001709.) Petitioner respectfully submits that the Examiner did not fully appreciate Coffman's interrelated teachings.

As discussed below, a step of "determin[ing] [by one or more physical processors at a vehicle] whether the command or query is to be executed on-board or off-board the vehicle" is taught and rendered obvious by Coffman. This is because, as detailed below, Coffman discloses a conversational computing system

having a distributed architecture involving devices that are both on-board and off-board a vehicle. (Coffman, e.g., Fig. 7, 55:12-21, 55:27-30, 56:16-57:9.) The “heart” of such a distributed system is a “conversational virtual machine” (CVM) “which acts as the conversational arbitrator between a plurality of applications 706, devices 708-713, other CVM applications or devices 707 and conversational resources 705.” (Coffman, 55:12-16; *see also id.*, 23:29-31 (devices “execute commands transmitted from the CVM.”).) The CVM, Coffman explains, “may be located on a server, a PC, etc”—and in one detailed illustrative use case, the CVM is part of “a conversational car PC (PVA, Personal Vehicle Assistant)” located on-board the user’s car. (Coffman, 55:13-14, 56:16-25 (“PVA runs an embedded CVM”).) The PVA includes its own functionality, such as vehicle navigation, and also acts as a “central unit,” such that “[w]hen a user interacts with the collection of devices, the interaction may always be via [the PVA].” (Coffman, 55:27-28, 57:5-9 (navigation).) The “central unit,” Coffman explains, “will parse and dispatch commands to each device accordingly.” (Coffman, 55:27-30.) Thus, the user can control other devices such as a conversational PDA and conversational smart phone “through the CVM running on the PVA,” as well as control the user’s desktop PC located off-board the vehicle. (Coffman, 56:22-57:9, 58:3-11, 56:22 (“It does not matter if your desktop PC runs a CVM”).) Therefore, and as more fully explained below, Coffman teaches and renders obvious this claimed “determining” step.

Additionally, the proposed prior art ground points to **Julia** and **Cooper**, both of which further confirm the obviousness of executing commands both on-board and off-board the vehicle—in Julia’s case, commands for an in-vehicle navigation application, and in Cooper’s case, commands for an off-board desktop e-mail application. The combination, as explained in the analysis below, therefore discloses and renders obvious Coffman’s PVA determining whether a command should be executed on-board or off-board the vehicle. Cooper was not cited during prosecution. Julia was listed among twenty-six other references on an applicant-submitted IDS (EX1019, pp.001245-001246) (after approximately five hundred other references had already been disclosed (*see* EX1019, pp.00139-00150, 00872-00877, 001087-001089)), but it was not otherwise cited or discussed.

IV. OVERVIEW OF THE CHALLENGED PATENT

A. Person of Ordinary Skill in the Art (POSITA)

A POSITA as of July 15, 2002² would have possessed a bachelor’s degree in electrical engineering or computer science, and two years of work experience in the field of speech- and natural language-interfaces for interacting with and directing computer systems and software, including speech recognition and processing (and

² Petitioner does not concede a July 15, 2002 priority date, but that issue is irrelevant because the prior art here predates that date.

more experience could substitute for less formal education, and vice versa).
(EX1002 (“Schmandt”), ¶¶11-15.)

B. Specification Overview

The ’845 patent states that it “relates to the retrieval of online information or processing of commands through a speech interface in a vehicle environment.”
(’845, 1:16-18; *see also id.*, 1:60-65 (acknowledging technology background).)

Figure 5 illustrates an embodiment:

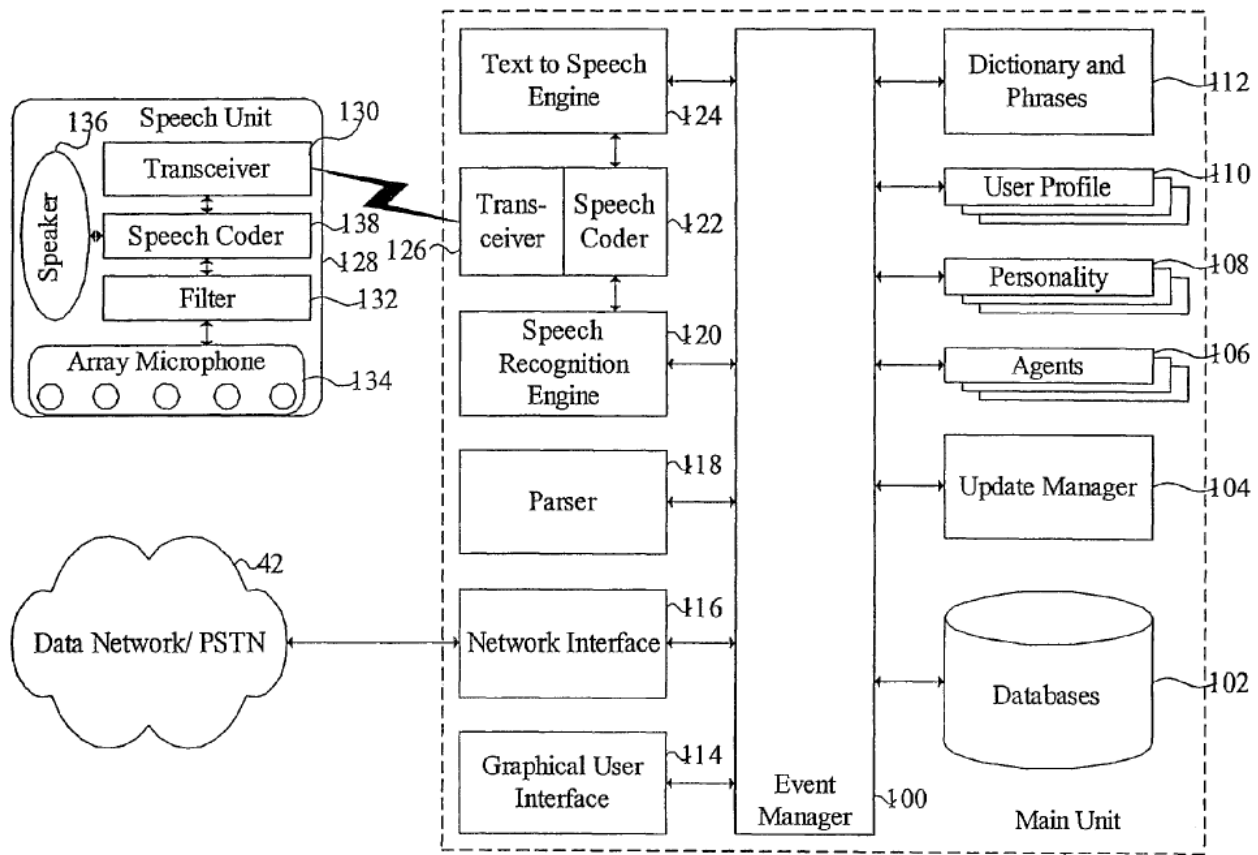


Figure 5. Speech Processing System Block Diagram

(’845, Fig. 5; *see also id.*, 15:47-49.) As shown, the system includes speech unit
128, speech recognition engine 120, dictionary and phrases 112, parser 118, and

agents **106**. Speech unit **128** includes a microphone to receive a spoken utterance. ('845, 20:29-31.) After initial processing, speech is passed to speech recognition engine **120** for processing “using data in the dictionary and phrases module **112** and received from the agents **106**.” ('845, 20:65-67.) Then, “[t]he recognized words and phrases may be processed by the parser **118**, which transforms them into complete commands and questions using data supplied by the agents.” ('845, 21:1-3.) To select the correct agent to process the command or question, “[t]he parser may determine a context” and “[o]nce the context for the question or command has been determined, the parser **118** can invoke the correct agent **156, 150**.” ('845, 27:60-63, 28:25-27.)

V. CLAIM CONSTRUCTION

Express claim construction is unnecessary because the prior art renders the challenged claims obvious under any reasonable construction. Petitioner reserves its right to respond to Patent Owner’s claim construction arguments.

VI. THE CHALLENGED CLAIMS ARE UNPATENTABLE

A. Overview of Grounds

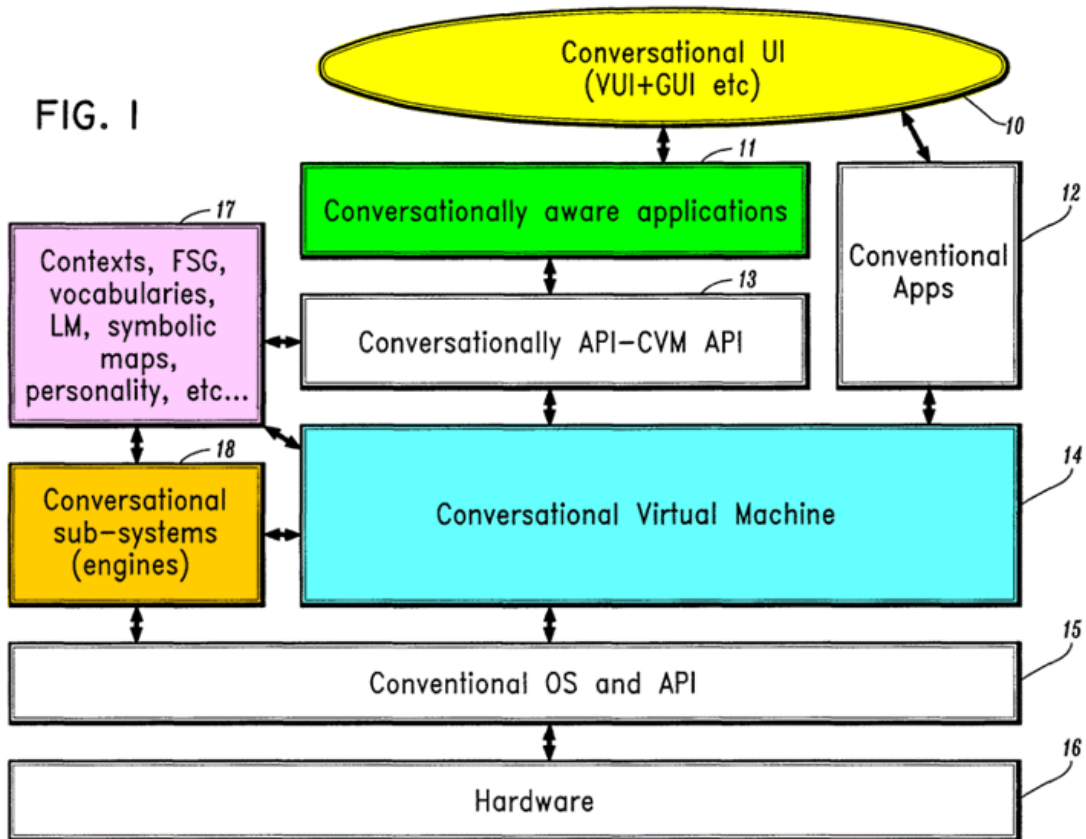
The challenged claims attempt to lay claim to well-known techniques for speech-enabled interfaces for interacting with and directing computer systems. The claims employ standard techniques that were known long before the '845 patent,

such as using speech recognition and natural language processing to select an application for carrying out a query and/or command included in a spoken utterance.

B. Brief Summary of Prior Art

1. Coffman (EX1003)

Coffman, International Patent Application Publication No. WO 00/20962, is entitled “Conversational Computing via Conversational Virtual Machine.” Coffman describes a “conversational computing system,” which generally speaking, is a computing system with either purely voice input/output or a multi-modal system with voice input/output (and other input/output means). (Schmandt, ¶42 (citing Coffman 31:24-25).) Coffman’s system “provides the capability to have a natural dialog with NLU [natural language understanding], NLG [natural language generation] and mixed initiative across multiple applications, with multiple domains.” (Coffman, 36:29-31.) Figure 1 provides a high-level depiction of the system’s components:



(Coffman, Fig. 1 (highlighting added).) As shown, the system includes, among other things, “a universal coordinated multi-modal conversational user interface (CUI) (10),” “a plurality of conversationally-aware applications (11),” “conversational kernel (14),” “conversational arguments (17),” and “conversational engines (18).” (Coffman, face page (Abstract).)

The conversational virtual machine (or CVM) is “[a] key component” of the conversational system that “allows a computer or any other interactive device to converse with a user.” (Coffman, 10:23-25, 11:10-13.) To do this, the CVM uses conversational subsystems “including speech recognition, speaker recognition, text-to-speech, natural language understanding and natural dialog generation engines to

understand and generate dialog between user and machine.” (Coffman, 5:1-4.) The conversational subsystems “are responsible for converting voice requests into queries and converting outputs and results into spoken messages using the appropriate data files 17 (e.g., contexts, finite state grammars, vocabularies, language models, symbolic query maps etc.).” (Coffman, 13:8-11.)

“The conversationally aware applications 11 are applications that are specifically programmed for operating with a CVM core layer (or kernel) 14 via conversational application APIs 13.” (Coffman, 12:28-30.) A conversationally aware application operates with a CVM by providing its data files to the CVM for use by the conversational subsystems—an application hooks into the CVM API by “registering all its conversational capabilities and requirements” including “[a]ctive vocabulary, finite state grammar and language models to control the application” and “[s]ymbolic mapping if NLU is supported or at list [sic; least] a context state list.” (Coffman, 57:23-27; *see also id.*, 36:31-32 (“each application will provide the CVM 401 its own parsing and translation arguments”).)

The CVM uses a “context stack,” which enables the system to determine the appropriate conversationally-aware application for handling user input. (Coffman, 42:16-17 (“This stack structure allows to maintain non-ambiguous conversational interactions with multiple tasks, threads or processes.”), 11:10-13.) The context stack “comprises all the information associated with an application.” (Coffman,

21:23-26; *see also id.*, 29:20-32.) To identify the appropriate conversationally-aware application to handle a user's input, the system searches the context stack for a context that is an acceptable match. (Coffman, 42:1-5, 37:2-4, 42:18-19.)

Coffman describes an embodiment of its conversational computing system having a distributed architecture where, “[w]hen a user interacts with the collection of devices, the interaction may always be via a central unit such as a PVA (personal vehicle assistant) 710 in a car[.]”³ (Coffman, 55:27-28.)

2. **Julia (EX1004)**

Julia, U.S. Patent No. 7,036,128 B1, is entitled “Using a Community of Distributed Electronic Agents to Support A Highly Mobile, Ambient Computing Environment.” Julia discloses a computing “[a]rchitecture to provide a human-machine interface in a car environment.” (Julia, 5:1-3.) Like Coffman's in-vehicle PVA conversational computing system, Julia describes how, “[u]tilizing speech, 2D and 3D gesture, and natural language recognition and understanding, the interface allows the driver to interact with the navigation system, control electronic devices, and communicate with the rest of the world.” (Julia, 5:3-7; *see also id.*, 30:25-45.)

The Petition cites Julia primarily for its disclosure of an express example of receiving a natural language utterance for interacting with a navigation application in connection with 1[b], and for claim 8 (and similar limitations in other claims).

³ All underlining added unless noted.

Referring to Figure 17, shown below, Julia discloses a voice-enabled in-car interactive navigation application:

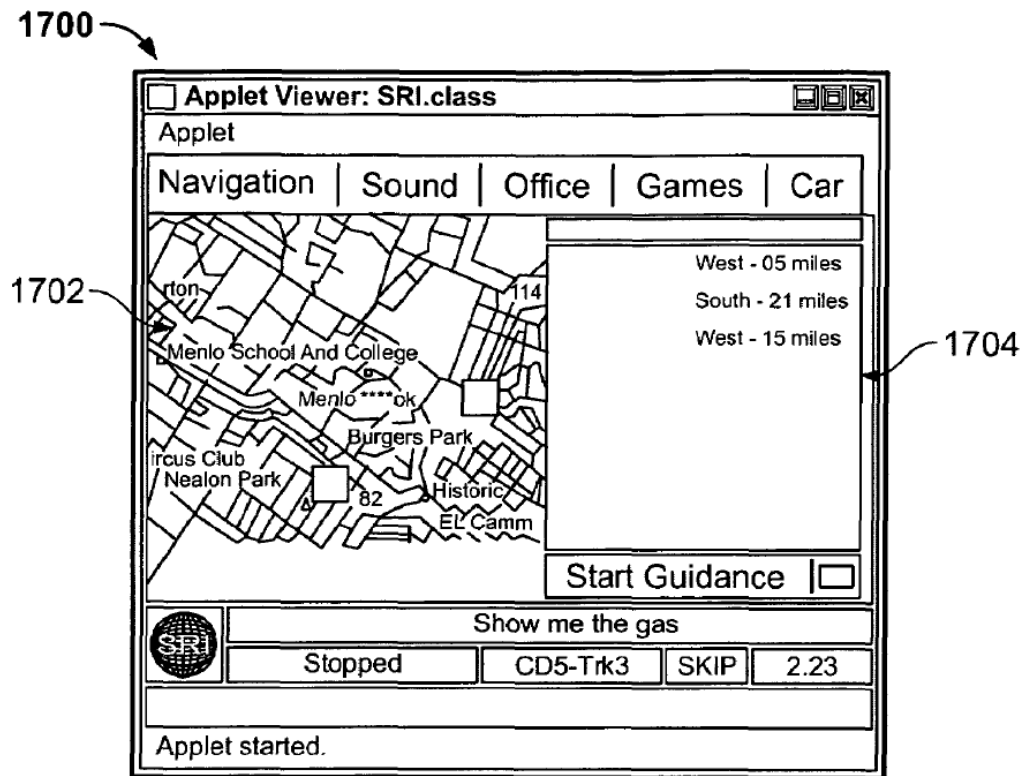


FIG. 17

(Julia, Fig. 17; 31:17-24.) Julia explains that, “[f]or example, in response to a user voice command such as ‘I want to go to SRI on Ravenswood Avenue in Menlo Park,’ one or more GPS agents, map database agents (and/or online search agents), and multimodal map interface agents would preferably collaborate to ascertain the user’s current location and display a suggested route from there to the desired destination and/or by speaking the suggested directions [].” (Julia, 31:24-31; *see also id.*, 31:31-35.)

3. Cooper (EX1005)

Cooper, U.S. Patent No. 6,757,362 B1, is entitled “Personal Virtual Assistant.” Cooper describes a “virtual assistant” computer system that “allows the mobile professional to access personal, company, and public information, including contacts, schedules, and databases from any interactive device, such as telephone.” (Cooper, 1:27-30.) Similar to Julia, the Petition cites Cooper primarily for its disclosure of an express example of a receiving natural language utterance for interacting with an e-mail application in connection with 1[b] and a “search query” as recited in independent claim 32[f].

Like Coffman, Cooper describes a voice-enabled email application accessible over a communications network. (Cooper, 6:55-58.) Cooper picks up where Coffman leaves off, however, expressly describing specific examples of accessing e-mail using natural language utterances such as “read my e-mail messages from John Smith.” (Cooper, 50:33-40, 52:1-18.)

C. Ground 1: Claims 1-3, 7-8, 10, 12, 15, 17-20, 24-25, 27, 29, 32, 34, 36 Are Obvious Over Coffman In View Of Julia And Cooper.

1. Claim 1: “A mobile system for processing natural language utterances, comprising:” (1[pre])

As discussed below, Coffman discloses a “**mobile system**” comprising a conversational computing system involving a “conversational car PC (PVA, Personal Vehicle Assistant)” running a “conversational virtual machine” (CVM)

with which a user can interact via natural language speech utterances. (Schmandt, ¶58; Coffman, 56:21-25 (“[Y]ou have, in your car, a conversational PDA, a conversational car PC (PVA, Personal Vehicle Assistant) and a conversational smart phone. Further assume that the PVA runs an embedded CVM while the two other applications are conversationally aware, i.e., you can control them through the CVM running on the PVA.”), 55:27-28.) As shown below, the prior art discloses and renders obvious the claimed system.

- (a) **“one or more physical processors at a vehicle that are programmed to execute one or more computer program instructions which, when executed, cause the one or more physical processors to:”**

A POSITA would have understood and found it obvious for an in-vehicle PVA in Coffman’s conversational computing system to comprise **“one or more physical processors,”** which describes nothing more than routine and conventional components and operation of a generic computer system. (Schmandt, ¶59.) Such a system operates by using physical processor(s) (e.g., a microprocessor) to execute computer program instructions. (*Id.* (citing EX1009, pp.0024-0025 (“microprocessor”), p.0033 (“software”), p.0028 (“program”), p.0016 (“execute,” “executable program”))).) Coffman explains that its “CVM system may be built on top of a conventional OS and APIs 15 and conventional device hardware 16 and located on a server or any client device (PC, PDA, PvC).” (Coffman, 13:1-3.) A POSITA, therefore, would have understood and found it obvious that a PVA for

processing natural language inputs in Coffman's conversational computing system would comprise **“one or more physical processors at a vehicle”** and for the one or more processors to be **“programmed to execute one or more computer program instructions”** to perform the claimed steps, which describe operations performed by a combination of computer hardware and/or software. (*Id.*)

(b) **“receive a natural language utterance associated with a user;”**

Coffman describes receiving spoken input from a person using, for example, a telephone or microphone, as Figure 6 illustrates:

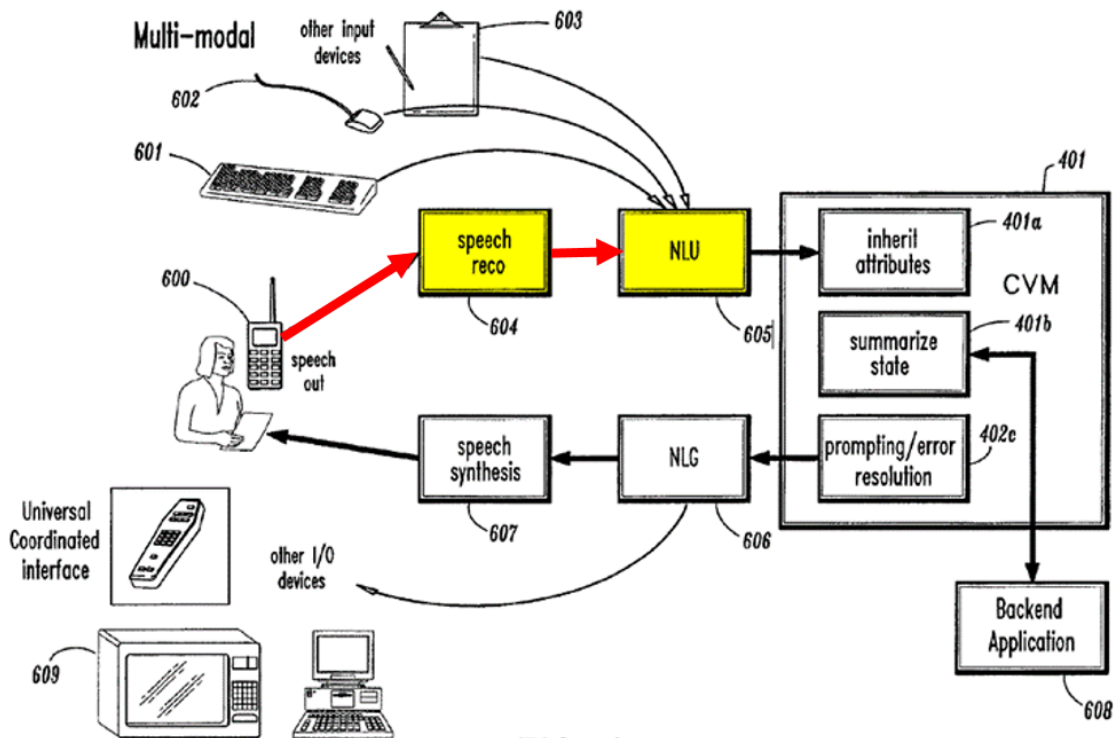


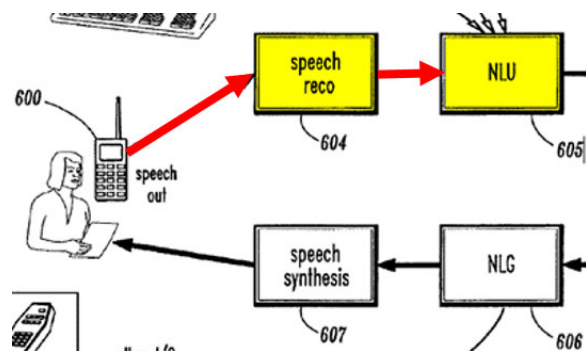
FIG. 6

(Coffman, Fig. 6 (annotations added), 53:8-10, 37:14-22, 41:2-5, 53:22-26.)

As to the PVA in particular, Coffman specifically describes how it can receive spoken input from a user. (Coffman, 55:27-28 (“When a user interacts with the collection of devices, the interaction may always be via a central unit such as a PVA (personal vehicle assistant) 710 in a car....”), 56:26 (“You can instruct the PVA to dial in your PC using the phone.”), 57:5 (“You now instruct the PVA to guide you turn by turn to the store”).)

Coffman further discloses receiving spoken input that is a “**natural language utterance**.” In the context of the ’845 patent, and speech recognition and processing generally, “**natural language**” generally speaking refers to the forms of language used by a person for communication (e.g., English), as opposed to artificial or formal languages such as programming languages. (Schmandt, ¶62 (citing EX1008, p.0035 (defining “natural language”))). Consistent with this understanding, Coffman discloses several examples of natural language speech utterances, such as “I would like to open this,” “it should be synchronized with my PDA,” and “e-mail from my mother.” (*Id.*; Coffman, 6:32-7:3, 48:1-3, 56:25-57:2.)

Coffman’s receipt (and use) of spoken “**natural language**” input is also confirmed by Coffman explaining that the spoken input, once recognized, is subjected to natural language parsing and understanding (“NLU”) processes, as



illustrated by Figure 6 (excerpted and annotated at right). (Coffman, e.g., Fig. 6, 53:22-26 (“User inputs can be represented by their transcriptions or their mappings into a symbolic language after parsing and NLU.”), 53:8-10 (“Voice input from a microphone (e.g. the standard voice input) can be arguments of function calls with the sequence of words, phones, or user identity or queries (symbolic language representation provided by NLU.”), 37:14-22, 5:1-4, 37:29-32.)

Numerous other examples of “**natural language utterances**” would have been obvious. (Schmandt, ¶64.) For example, Coffman discloses using applications for e-mail, spreadsheets, vehicle navigation, addressbook, name dialing, and news/stocks, as well as a “conversational search capability,” with its conversational system. (*Id.*; Coffman, e.g., 16:35-17:7 (listing e-mail, calendar, addressbook), 45:2-3 (name dialer application), 56:11-15 (spreadsheet), 57:5-9 (vehicle navigation), 57:17-20 (e-mail, news, stock quotes, spreadsheet), 58:30-32 (spreadsheet), 8:24-27 (conversational search).) With respect to vehicle navigation, for instance, Coffman describes how “[y]ou now instruct the PVA to guide you turn by turn to the [grocery] store.” (Coffman, 57:5-7.) For e-mail, Coffman describes how a user can access their e-mail application remotely and “drive home and while driving quietly listen to your e-mail....” (Coffman, 57:17-18, 56:11-12.) Similarly, for spreadsheets, Coffman describes how “a spreadsheet, written for voice access over the desktop, can now be accessed by voice over the phone by relying on the

phone capabilities.” (Coffman, 58:30-32, 57:18-20.) A few of these example applications are addressed in more detail next.

Vehicle Navigation Application

It would have been obvious for the system to receive a natural language utterance from the user such as “driving directions to 100 Main Street” or “navigate to Market Basket” to instruct the vehicle navigation application to provide turn by turn driving directions. (Schmandt, ¶65.) This is confirmed by **Julia** (EX1004). (*Id.*) Referring to Figure 17, shown below, Julia describes an in-car interactive navigation application that can receive a user voice command for driving directions and “take advantage of GPS information to plan a route and get location-specific information from local or remote databases displayed on the map.” (Julia, 31:18-24; *see also id.*, 5:3-7, 30:25-45.)

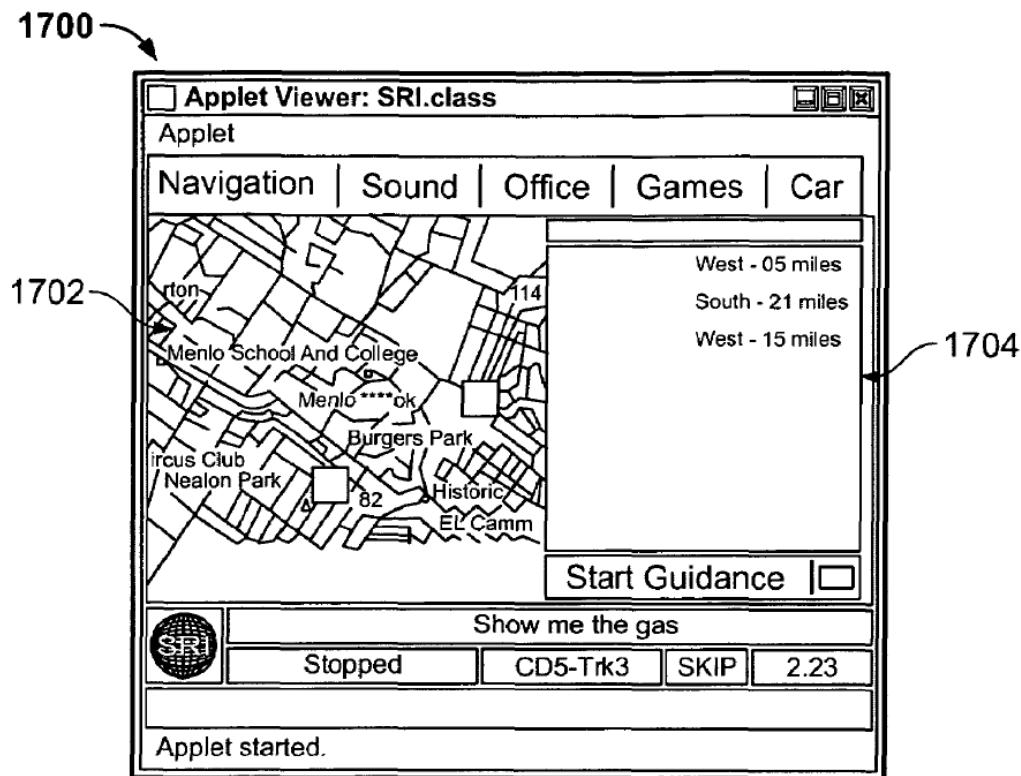


FIG. 17

(Julia, Fig. 17.) Julia explains that, “[f]or example, in response to a user voice command such as ‘I want to go to SRI on Ravenswood Avenue in Menlo Park,’ one or more GPS agents, map database agents (and/or online search agents), and multimodal map interface agents would preferably collaborate to ascertain the user’s current location and display a suggested route from there to the desired destination. and/or by speaking the suggested directions [.]” (Julia, 31:24-31.) Julia therefore discloses a specific example of receiving a natural language utterance from a user—“I want to go to SRI on Ravenswood Avenue in Menlo Park”—in the context of a vehicle navigation application. (Schmandt, ¶65.)

Rationale and Motivation to Combine (Coffman and Julia). It would have been obvious to combine Julia’s teachings regarding a vehicle navigation application with Coffman’s PVA and conversational system. (Schmandt, ¶66.) It would have involved combining known prior art elements and techniques according to known methods and would have predictably resulted in receiving a natural language utterance from a user to interact with Coffman’s navigation application. (*Id.*) The combination therefore discloses a specific example of “**receiv[ing] a natural language utterance,**” as claimed. (*Id.*)

Coffman and Julia are analogous references in the same field as the ’845 patent and would have been reasonably pertinent to problems facing the ’845 patent inventors, including how to interact with speech-enabled applications in a vehicle environment. (Schmandt, ¶67; ’845, 1:16-25; Coffman, e.g., Abstract; Julia, 5:1-7, 30:25-45.)

A POSITA would have been motivated to combine. (Schmandt, ¶68.) Coffman provides only a brief description of how the user interacts with the navigation application, and a POSITA would have appreciated that Julia beneficially provides additional detail for implementing or adapting the application. (*Id.*) As discussed, Coffman discloses an in-vehicle navigation application that the user can instruct to provide driving directions to a particular destination, and as a conversational system, a POSITA would have understood that the user’s instruction

would be provided in the form of a spoken utterance identifying the destination—particularly in a vehicle environment, where hands-free and eyes-free interaction was recognized as especially beneficial. (Schmandt, ¶68 (citing Coffman, 9:12-25 (“conduct efficiently task in hand-free and/or eye-free mode”); EX1009, p.007 (“Auto PC”), p.0018 (“hands-free kit”))).) A POSITA would have readily appreciated that Julia, describing a similar in-vehicle speech-enabled navigation application, picks up where Coffman leaves off, providing further detail regarding how a user provides spoken input for the navigation application, and how that spoken input is used. (Schmandt, ¶68 (citing Julia, 38:18-25).) A POSITA would have therefore recognized that Julia’s techniques would improve Coffman’s similar system in the same way. (*Id.*)

A POSITA would have had a reasonable expectation of success, as the combination would have involved the application of known technologies and techniques, without undue experimentation. (Schmandt, ¶69.)

E-mail Application

Coffman generally describes how a user can access their e-mail application remotely and “drive home and while driving quietly listen to your e-mail...” but does not detail a specific natural language utterance from a user. (Coffman, 57:17-18, 56:11-12; Schmandt, ¶70.) Consistent with Coffman’s teachings regarding making ordinary application functionality accessible by voice, a POSITA would

have found it obvious for Coffman’s system to enable a user to access their e-mail application by voice—for example, to search for particular e-mail messages and listen to them by providing a natural language utterance such as “read my e-mail messages from John Smith.” (Schmandt, ¶70.) This is confirmed by **Cooper (EX1005)**. (*Id.*) Like Coffman, Cooper describes a voice-enabled email application accessible over a communications network. (Cooper, 6:55-58, 1:27-30.) Cooper, in fact, expressly discloses accessing e-mail using the natural language utterance “read my e-mail messages from John Smith.” (Cooper, 52:1-18 (“Read my e-mail messages from John Smith”), 50:33-40.)

Rationale and Motivation to Combine (Coffman and Cooper). Combining Cooper’s teachings regarding a voice-enabled email application with Coffman’s PVA and conversational system would have been obvious. (Schmandt, ¶71.) It would have involved combining known prior art elements and techniques according to known methods and would have predictably resulted in receiving a natural language utterance from a user at the PVA to interact with an off-board email application. (*Id.*) Like Julia, the combination therefore discloses another specific example of “**receiv[ing] a natural language utterance,**” as claimed. (*Id.*)

Cooper is an analogous reference in the same field as the ’845 patent and would have been reasonably pertinent to problems facing the ’845 patent inventors,

including how to interact with speech-enabled applications in a vehicle environment.
(Schmandt, ¶72; '845, 1:16-25; Cooper, 1:27-31.)

A POSITA would have been motivated to combine. (Schmandt, ¶73.)
Coffman provides only brief mentions of a user interacting with e-mail, and a POSITA would have appreciated that Cooper beneficially provides additional detail for implementing or adapting an e-mail application. (*Id.*) As discussed above, Coffman already discloses that its conversational system enables a user to interact with an e-mail application, and a POSITA would have understood that the user would interact with the application via the conversational system in the form of a natural language utterance. (*Id.*) A POSITA would have readily appreciated that Cooper, describing a similar e-mail application accessible by voice from a mobile environment, picks up where Coffman leaves off, providing further detail regarding the natural language utterances that a user could provide to interact with the application. (*Id.* (citing Cooper, 50:33-40, 52:1-18).) The example natural language utterances are for carrying out well-known, conventional e-mail operations, and as such a POSITA would have readily appreciated that they would beneficially enable a user to access their e-mail in a known and expected manner. (*Id.* (citing Cooper, 52:1-18).) A POSITA would have therefore recognized that Cooper's teachings would improve Coffman's similar system in the same way. (*Id.*)

A POSITA would have had a reasonable expectation of success, as the combination would have involved the application of known technologies and techniques without undue experimentation. (Schmandt, ¶74.)

Spreadsheet Application

Consistent with ordinary use of a spreadsheet application, a POSITA would have understood and found it obvious that Coffman’s descriptions of “consulting or modifying some spreadsheet information” via “voice over the phone by relying on the phone capabilities” (Coffman, 58:30-32, 57:18-20) would involve receiving a natural language utterance from a user. (Schmandt, ¶75.) For example, to locate a particular field or cell in a spreadsheet, it would have been obvious for the user to provide a spoken query that included words or phrases associated with the field or cell. (*Id.*) Even casual users of spreadsheet software such as Microsoft Excel can easily envision these sorts of spoken commands, such as “what is the value of the bottom row of the ‘profit’ column” or any other voice-enabled lookup function. (*Id.* (citing EX1009, p.0023 (defining “lookup”)).)

Search Application

As one final example, Coffman as noted describes a “conversational search capability.” (Coffman, 8:24-27.) Coffman explains that, as with other conversationally aware applications, with conversational search, “search queries can be provided in a natural fashion,” and “decoded, parsed and then translated into a

logic combination of queries (symbolic query) using NLU technology.” (Coffman, 51:8-14.) “Conventional structures as well as categories and dynamic information can then be searched to match the symbolic query.” (Coffman, 51:14-15.) A POSITA therefore would have understood that Coffman’s PVA could also receive a natural language utterance in connection with the conversational search capability. (Schmandt, ¶76.)

Lastly, a POSITA would have understood that, because the PVA receives spoken natural language input from a person, the PVA receives “**a natural language utterance associated with a user**.” (Schmandt, ¶77.) Moreover, Coffman also discloses identifying the user who spoke the utterance, thereby associating the utterance with a user. (*Id.*) Coffman explains that its conversational system includes functionality for “speaker recognition” and “speaker identification,” such as “text-independent speaker identification” using a user’s command or request. (*Id.* (citing 5:1-4, 47:25-27, 8:16-18; EX1008, pp.0047 (“speaker recognition”), 0051 (“text-independent speaker recognition system”))).) Figure 4 of Coffman depicts conversational system 400 with speaker recognition system 409:

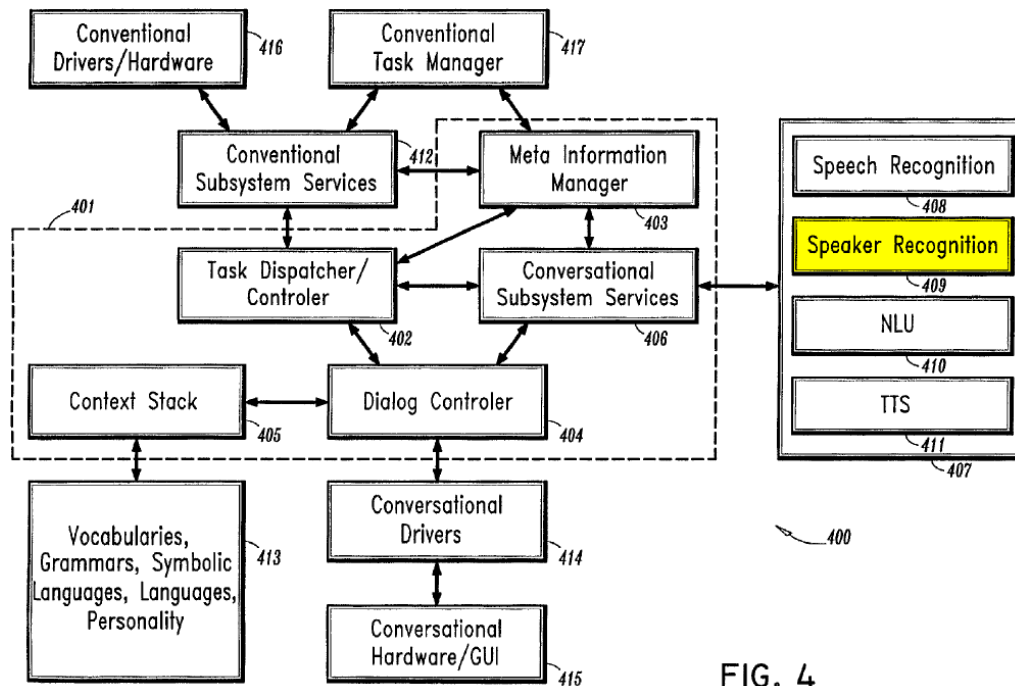


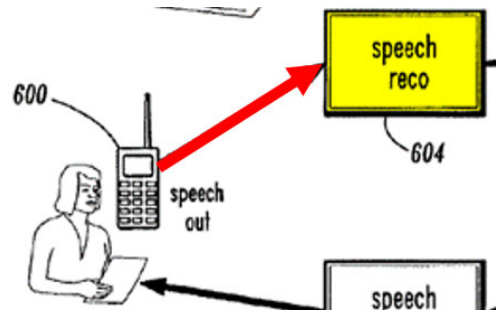
FIG. 4

(Coffman, Fig. 4 (highlighting added); *see also id.*, 29:1-4, 15:13-16:24 (listing various speaker-identification-related “conversational foundation classes” including “Acoustic Identify speaker,” “Acoustic Verify speaker” and “Speech biometrics identify.”).) Coffman’s speaker identification is therefore another way that Coffman discloses “**a natural language utterance associated with a user**.” (*Id.*)

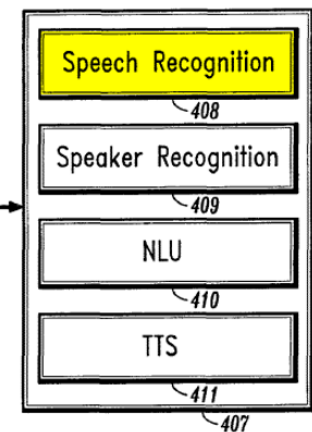
(c) **“perform speech recognition on the natural language utterance;”**

As noted for 1[b], Coffman discloses **“perform[ing] speech recognition on the natural language utterance.”** For example,

Coffman explains how “[s]poken input from a speech client (e.g., telephone 600) is subject to a speech recognition process 604.” (37:14-22; *see*



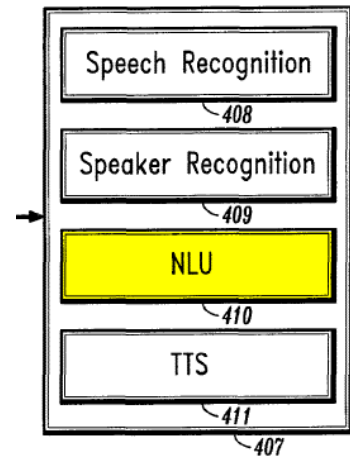
also id., Fig. 6 (step 604; excerpted and annotated at right); 38:9-11 (“Each input stream is converted into its ASCII transcription and aligned with input time- marks via speech recognition processing 604.”).) Coffman explains that the speech recognition is performed by a “speech recognition engine.” (illustrated at right (Fig. 4 (excerpt))). (Coffman, 10:28-30 (“speech recognition engine”), 29:1-4 (“speech recognition system 408”), Fig. 4.)



(d) **“parse and interpret the speech recognized natural language utterance;”**

Coffman teaches that the user’s recognized speech (**“the speech recognized natural language utterance”**) is **“pars[ed] and interpret[ed]”** in connection with performing natural language understanding to determine a command/query to execute. (Schmandt, ¶79.)

Coffman conversational system includes “a natural language understanding and natural language parsing system [NLU] 410” (illustrated at right (Fig. 4 (excerpt))). (Coffman, 29:1-4, Fig. 4.) Coffman explains that, “using NLU technology,” “[q]ueries can be decoded, parsed and then translated into a logic combination of queries (symbolic query).” (Coffman, 51:12-14; *see also id.*, 40:12-15 (describing “parsers that must operate on a query for the dialog controller 404 to identify the active context and complete the query”).) Coffman discloses that the parsing and translating process uses data files (such as grammars and symbolic query maps) provided by registered applications. (Coffman, e.g., 24:31-25:9, 57:22-27, 36:31-32 (“each application will provide the CVM 401 its own parsing and translation arguments”); Schmandt, ¶80.) For example, Coffman explains that these data files are used to interpret the recognized utterance by matching the utterance against them, describing performing a search in order to “fit[] the utterance at the level of the NLU symbolic language.” (Coffman, 42:1-7; Schmandt, ¶80.) As a result of the search, Coffman explains, “[t]he first task to have its dialog completed by producing a non-ambiguous query is executed, and the corresponding query as interpreted by the other application is stored to activate if the recognized query is rejected by the user.” (Coffman, 37:2-4.)



To take a more concrete example, Coffman describes how a user's utterance "I would like to open this" is "disambiguated into a set of attributes value pairs: Command: Open, Object: Window or task selected by the last mouse click." (Coffman, 6:32-7:3.) Based on Coffman's teachings about natural language parsing and understanding, a POSITA would have understood that the "disambiguation" process would involve "**pars**[ing]" the utterance—for example, identifying words and phrases in the utterance—and then "**interpret**[ing]" the utterance by performing a search to match the parsed utterance—for example, by matching the word "open"—against application data files to identify the appropriate application and command to process the request. (Schmandt, ¶81 (citing EX1009, p.0027 ("parse"); Coffman, 52:12-14 ("Queries can be decoded, parsed and then translated into a logic combination of queries (symbolic query) using NLU technology."))).

(e) **"determine a domain and a context that are associated with the parsed and interpreted natural language utterance;"**

As explained below, Coffman discloses "**determin**[ing] **a domain and a context**" as claimed in connection with identifying a context entry and associated application as a result of the process of matching the parsed utterance as described for 1[d]. (Schmandt, ¶82.)

In particular, Coffman explains that the conversational system includes a "context stack" where the "context," including the state and history of each active

process, such as an application process, is tracked. (Coffman, 29:18-30:3, 30:4-10.)

Coffman further explains that the process of matching the parsed utterance involves searching context entries in the context stack, and that the context entry that is an acceptable match is activated, as is the application associated with that context entry. (Coffman, 42:1-5; *see also id.*, 41:2-5, 30:8-10, 51:12-14, 40:12-15, 37:20-26.)

Coffman therefore “**determines a context**” that is “**associated with the parsed and interpreted natural language utterance**” by identifying and activating a “**context**” in a context stack that matches the user’s parsed utterance (“**the parsed and interpreted natural language utterance**”). (Schmandt, ¶¶83-85.) The activated context is “**associated with**” the parsed utterance because it contains information that matches the utterance, and because it is selected for processing the utterance. (Schmandt, ¶85.)

With respect to determining a “**domain**,” to summarize the analysis below, Coffman teaches that applications provide data to the system that is used for input understanding and in determining which application should process a user’s request, such as vocabulary, symbolic query maps and possible queries. (Schmandt, ¶86.) Because the applications address different areas of functionality, i.e. different domains—such as email, calendar, contacts, and navigation—a POSITA would have understood and found it obvious that the data provided by the applications would comprise domain-specific information, such as domain-specific words and queries.

(*Id.*) And therefore, by using that data in determining that an application should process a user's request, a “**domain**” would be determined, as claimed. (*Id.*)

Turning to the analysis, as discussed, Coffman teaches that when a context is matched the associated application is activated. (Coffman, e.g., 42:1-5 (“As soon a context is found that fits the utterance at the level of the NLU symbolic language, the context becomes active and the corresponding process becomes active.”), 30:4-10.) An application in Coffman (referred to as a “conversationally aware application” (Coffman, 12:26-30)) is associated with a “**domain**” because it is associated with a specific area of functionality, such as subject matter or particular device. (Schmandt, ¶87.) For example, as noted for 1[b], Coffman describes applications for email, calendar, spreadsheets, contacts, vehicle navigation, search, and for devices, such as a PDA or smartphone. (*Id.* (citing Coffman, 16:35-17:7, 45:2-4, 56:11-14, 57:5-9, 58:30-32, 8:24-27, 56:22-25 (“conversational PDA,” “conversational car PC (PVA, Personal Vehicle Assistant),” and “conversational smart phone”), 36:29-31 (“[T]he present invention provides the capability to have a natural dialog with NLU, NLG and mixed initiative across multiple applications, with multiple domains.”))).) This understanding of “**domain**” is consistent with the written description of the '845 patent, which describes a “**domain**” broadly as an area of functionality. (Schmandt, ¶87 (citing '845, 4:31-33, 4:43-44, 4:46-48, 4:48-50).)

Coffman further explains that each application provides its own data to the conversational system that is used in determining whether the application should be activated to process a user command/query. (Schmandt, ¶88 (citing Coffman, 36:21-32 (“each application will provide the CVM 401 its own parsing and translation arguments”), 57:23-27 (application will “register[] all its conversational capabilities and requirements” including “[a]ctive vocabulary, finite state grammar and language models to control the application” and “[s]ymbolic mapping if NLU is supported or at list [sic; least] a context state list”))).) And Coffman makes clear that when a context stack is searched to identify an acceptable context entry matching the user’s input, this data is used. (Schmandt, ¶88.) As noted, Coffman discloses that a context entry is associated with application data, such as vocabulary, symbolic query maps and possible queries. (Coffman, 5:15-18, 30:8-10, 13:8-13). Coffman explains that the process to identify a matching context involves finding a context that “fits the utterance at the level of the NLU symbolic language” and that a “symbolic language” associated with a context “maps the actual query.” (Coffman, 42:1-5, 30:8-10; *see also id.*, 51:12-14.)

Therefore, because applications are associated with various “**domains**” and the application-provided data is used in determining which application should be used to process a user’s command/query, a POSITA would have understood and found it obvious that the application-provided data would comprise domain-specific

information. (Schmandt, ¶89.) For example, it would have been obvious that data provided by an email application would include email-specific information, such as vocabulary, query maps and commands. (*Id.*) Without this domain-specific information, the conversational system could not properly determine which application should process a user's command/query. (*Id.*) In fact, Coffman confirms that the applications would provide domain-specific data, describing how developers can build their own applications to use with the conversational system, which can involve "modifying the data files...for a given application" and incorporating data files "dedicated to the task, domain or acoustic environment." (*Id.* (quoting Coffman, 58:12-26).) Therefore, a POSITA would have understood and found it obvious that, by using domain-specific information to determine which context entry and associated application should be used to process a user's command/request, Coffman's conversational system would be "**determin[ing] a domain...associated with the parsed and interpreted natural language utterance,**" as claimed. (*Id.*)

(f) **"formulate a command or query based on the domain and the context;"**

Coffman discloses "**formulat[ing] a command or query based on the domain and the context**" by explaining that an application call ("**a command or query**") to execute is formulated based on the matched context and application. (Schmandt, ¶90.) In particular, as explained, Coffman describes translating a user's utterance into a "symbolic query" by searching the context stack to find a context

that “fits the utterance at the level of the NLU symbolic language” using parsing and translation arguments provided by the associated application. (*Id.* (citing Coffman, 51:12-14, 42:1-5, 36:31-32).) Coffman discloses that the symbolic query based on the context and application information is then transformed into an application call for the matched application to execute, explaining that voice requests are converted into queries and that, in turn, “the CVM...transform[s] queries into application calls.” (*Id.* (citing Coffman, 13:8-13, 58:9-11 (same), 22:5-7 (describing converting standardized attribute value n-uples into database queries), 24:4-9 (describing RMI (remote method invocation) and RPC (remote procedure call) to implement calls between applications and conversational engines)).)⁴ Coffman therefore teaches, for example, that a user’s utterance “I would like to open this” would be processed into “symbolic query” comprising “a set of attributes value pairs: Command: Open, Object: Window,” which would then be formulated into an “open” command for the associated application to execute. (*Id.* (citing Coffman, 6:32-7:3).) A POSITA would have understood that an application call would be “**a command or query**” because the call would instruct the system to perform an operation and/or retrieve

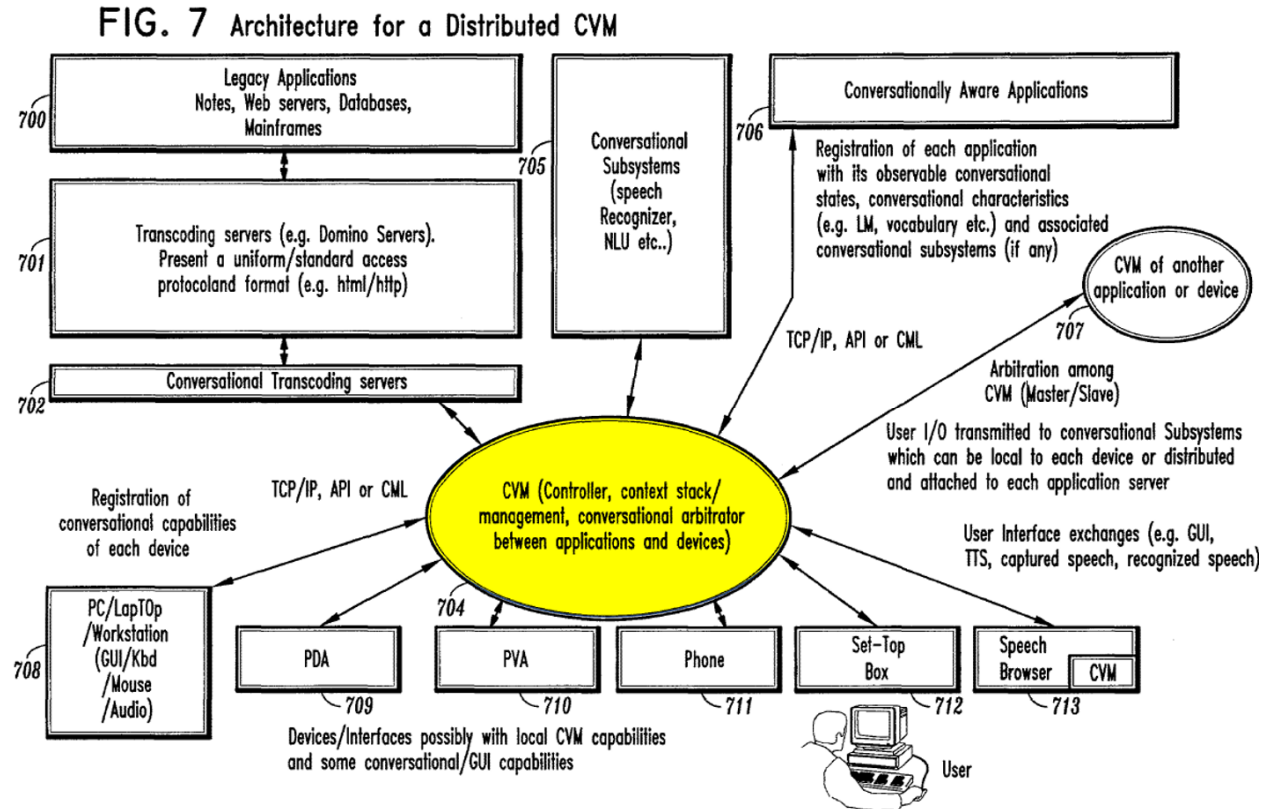
⁴ An “n-uple” refers to a collection of “n” elements; for example, a “4-uple” (or “quadruple”) is a collection of four elements, such as four attribute-value pairs. (Schmandt, ¶84 n.3.)

information, consistent with the conventional understanding of the functionality provided by an application call. (*Id.* (citing Coffman, 41:23-24 (“[C]omplete and non-ambiguous commands will result in certain results (e.g., outputs or actions).”)).)

(g) **“determine whether the command or query is to be executed on-board or off-board the vehicle;”**

As explained below, it would have been obvious for an in-vehicle PVA in Coffman’s conversational computing system to **“determine whether the command or query is to be executed on-board or off-board the vehicle.”** (Schmandt, ¶¶91-97.)

To start, Coffman discloses using a central device to determine where a command associated with a user’s request should be executed. (Schmandt, ¶92.) Coffman explains that the conversational system can have a distributed architecture involving multiple devices in communication with a central device that controls overall operation of the system. (Coffman, e.g., 55:12-21, 55:27-30.) Figure 7 of Coffman illustrates:



(Coffman, Fig. 7 (highlighting added).) Referring to Figure 7, Coffman explains that the user interacts with the central device, which acts as an arbitrator to determine which applications and devices are to process user requests. (Coffman, 55:12-21 (“The heart of the distributed system is a CVM 704 (which may be located on a server, a PC, etc) which acts as the conversational arbitrator between a plurality of applications 706, devices 708-713, other CVM applications or devices 707 and conversational resources 705.”); *see also id.*, 19:26-20:2 (“core CVM kernel layer 202” functionality includes “keeping track of which applications are registered (both local and network-distributed)”).) “When a user interacts with the collection of devices,” Coffman explains, “the interaction may always be via a central unit....The

task dispatcher and context stack accumulates the contexts associated to all the devices and will parse and dispatch commands to each device accordingly.”

(Coffman, 55:27-30.) The devices “execute commands transmitted from the CVM.”

(Coffman, 23:29-31.)

Coffman further discloses that the central unit in such a distributed system can take the form of an in-vehicle PVA and that the system involve devices both on-board and off-board the vehicle. (Schmandt, ¶93; Coffman, 55:27-28, Fig. 7 (PVA and other devices).) In fact, Coffman details an example distributed architecture involving a PVA acting as the central device, a PDA and a smartphone in a user’s car, and a desktop PC in another location. (Coffman, 56:16-25.) Coffman describes how, in this example architecture, the user interacts with the PVA, issuing various commands/queries that are executed by the devices, such as a command that the phone dial the desktop PC, commands to the desktop PC for “grocery list,” “it should be synchronized with my PDA,” and “Do it,” and a command to the PVA for turn-by-turn driving instructions to a store. (Coffman, 56:26-57:9; Schmandt, ¶93.) As discussed for 1[b], numerous commands based on natural language utterances from the user via the PVA would have been obvious, such as a command for driving directions based on the user’s utterance, “I want to go to SRI on Ravenswood Avenue in Menlo Park” as taught by Julia (Julia, 31:18-31), a command to listen to e-mail from a specific sender, such as “read my e-mail messages from John Smith,”

as taught by Cooper (Cooper, 50:33-40, 52:1-18), a command to consult or modify spreadsheet information (Coffman, 58:30-32, 57:18-20; EX1009, p.0023 (“lookup” function)), or a command to perform a conversational search (Coffman, 8:24-27, 51:8-15). (Schmandt, ¶93.)

A POSITA would have understood and found it obvious, based on Coffman’s teachings, that the PVA in the distributed architecture described above would be responsible for determining which device should execute the application call associated with a user request. (Schmandt, ¶94.) In fact, referring again to the example architecture, Coffman explains how the PVA acts as the “master CVM,” that among other things, “automatically handles all the input/output issues” and “transform[s] queries into application calls.” (Coffman, 58:3-11.) And therefore, in a situation where the architecture involved devices both on-board and off-board the vehicle, a POSITA would have recognized that, by determining which device should execute the application call, the PVA would be determining whether the application call (“**command or query**”) would be “**executed on-board or off-board the vehicle**” as claimed. (Schmandt, ¶94.)

For example, a POSITA would have understood and found it obvious, based on Coffman alone or in view of Julia’s teachings about an on-board navigation application, that a user’s utterance such as “I want to go to SRI on Ravenswood Avenue in Menlo Park” received by the PVA would result in the PVA determining

that the on-board navigation application should execute the associated application call. (Schmandt, ¶95.) Similarly, a POSITA would have understood and found it obvious, based on Coffman alone or in view of Cooper’s teachings about remotely accessing an e-mail application, that a user’s utterance such as “read my e-mail messages from John Smith” received by the PVA would result in the PVA determining that an off-board e-mail application should execute the associated application call. (*Id.*) The same is true for an off-board spreadsheet application or an on-board or off-board conversational search—such as a user’s command to locate and read back a “grocery list” stored on an on-board PDA. (*Id.* (citing Coffman, 57:2-4, 57:14-15).)

Petitioner notes that, in Coffman’s illustrative PVA use case, it is not clear whether the user’s desktop PC runs its own CVM. Coffman states, for example, that “[i]t does not matter if your desktop PC runs a CVM since you have, in your car...a conversational car PC (PVA, Personal Vehicle Assistant)” that “runs an embedded CVM.” (Coffman, 56:22-25.) However, whether the off-board desktop PC includes its own CVM or not, this limitation is disclosed and obvious over Coffman. (Schmandt, ¶96.) This is because, as an initial matter, Coffman’s PVA use case is merely illustrative. A POSITA would have understood and found it obvious that a distributed conversational computing system as taught by Coffman could comprise multiple devices on-board and off-board a vehicle, with only the vehicle’s PVA

running a CVM and issuing commands to the other devices. (*Id.* (citing Coffman, 24:10-18, 55:12-16, 23:16-17, 7:12-22 (all describing “conversationally aware” networked devices without their own CVM)).) There is no requirement in Coffman that, in a distributed system involving multiple devices, there must be more than one CVM. (*Id.* (citing Coffman, 11:2-4).)

Moreover, as noted above, Coffman makes clear that the PVA may act as the “master CVM” even if the desktop PC includes its own CVM. (Schmandt, ¶97 (citing Coffman, 58:3-11 (“When the PVA conversationally connects to the PC, it will be up to the application developer of the PVA, to decide if the PVA acts as master, slave or partner.”))).) A POSITA would have therefore understood and found it obvious that, as the master CVM, the PVA would use its own resources, such as speech recognition and context stack, to formulate a command/query to dispatch to the user’s desktop PC. (*Id.*) For example, Coffman explains that a CVM comprises a “dialog controller” (also referred to as a “dialog manager”). (Coffman, 28:24-27, 19:26-29.) The dialog controller manages the context stack and “manages and determines the active context and application.” (Schmandt, ¶97 (citing Coffman, 31:28-29, 29:18-19, 30:22-24).) In one embodiment involving a “master/slave configuration,” “only the main root (arbitrator) dialog manager performs the dialog manager task for one or more applications or devices (1, k, N).” (Coffman, 19:26-29, 27:28-30; *see also id.*, 27:30-32.) Therefore, even if the desktop PC included its

own CVM, it would have been obvious for the PVA, as “master CVM,” to formulate a command/query to dispatch for execution at an off-board device, such as the user’s desktop PC. (Schmandt, ¶97.)

(h) “execute the command or query at the vehicle in response to a determination that the command or query is to be executed on-board the vehicle; and”

As explained for 1[g], a POSITA would have understood and found it obvious for the PVA to determine the appropriate device and application for executing an application call and to accordingly have that device execute the application call. (Schmandt, ¶98.) For instance, as explained, Coffman teaches how a PVA, acting as the central device/master CVM, will transform a query into an application call and dispatch the application call to the appropriate device. (Coffman, e.g., 55:27-30 (“parse and dispatch commands to each device accordingly”), 58:9-11 (“Again, the API conveys all the information for the CVM to transform queries into application calls”).) The application call will then be executed at the device, and results may be returned. (Coffman, 37:2-3 (“The first task to have its dialog completed by producing a non-ambiguous query is executed”), 40:15-19 (“execute completed queries and update the context stack 405” and “return output of executed or running processes”), 56:16-17, 23:29-31 (devices “execute commands transmitted from the CVM.”).) Therefore, using Coffman’s PVA example discussed for 1[g] (Coffman, 56:16-57:21), a POSITA would have understood and found it obvious that the

application call for turn by turn navigation instructions would be executed at the PVA on-board the vehicle and, similarly, that the application call for the phone in the user's car to dial the user's desktop PC would be executed on-board the vehicle. (Schmandt, ¶98.) In both cases, by determining the appropriate device to handle the user's request, the application call would have been executed on-board the vehicle **“in response to a determination that the command or query is to be executed on-board the vehicle,”** as claimed. (*Id.*) Therefore, 1[h] is obvious. (*Id.*)

- (i) **“invoke a device that communicates wirelessly over a wide area network to process the command or query such that the command or query is executed off-board the vehicle in response to a determination that the command or query is to be executed off-board the vehicle.”**

As with 1[h], as explained, a POSITA would have understood and found it obvious for the PVA to determine the appropriate device and application for executing an application call and to accordingly have that device execute the application call. (Schmandt, ¶99 (citing Coffman, e.g., 55:27-30, 58:9-11, 37:2-3, 40:15-19, 56:16-17, 23:29-31).) In fact, Coffman describes well-known “RMI” (“remote method invocation”) and “RPC” (“remote procedure call”) techniques for a device to invoke another device to execute an application call. (*Id.* (citing Coffman, 24:4-9).) Therefore, using Coffman's PVA example discussed above for 1[g] (Coffman, 56:16-57:21), a POSITA would have understood and found it obvious that, using RMI, RPC or a similar technique, the application calls for

executing user commands such as “grocery list,” “it should be synchronized with my PDA” and for otherwise interacting with applications on the desktop PC would be executed at the desktop PC off-board the vehicle. (*Id.* (citing Coffman, 56:26-32; *see also id.*, 57:18-20).) For example, a POSITA would have understood and found it obvious that, in order for the user to interact with spreadsheet and e-mail applications at the desktop PC off-board the vehicle (Coffman, 56:11-12, 57:17-20, 58:30-32), the application calls for carrying out user commands would be executed at the desktop PC. (Schmandt, ¶99 (citing Coffman, e.g., 55:27-30 (“parse and dispatch commands to each device accordingly”), 23:29-31 (devices “execute commands transmitted from the CVM.”))).)

Therefore, like 1[h], Coffman teaches that by determining the appropriate device to handle the user’s request, the application call would have been executed off-board the vehicle **“in response to a determination that the command or query is to be executed off-board the vehicle,”** as claimed. (Schmandt, ¶100.)

Coffman further makes clear that, in order for the PVA to communicate with the desktop PC to execute an application call off-board the vehicle, the PVA would **“invoke a device that communicates wirelessly over a wide area network.”** (Schmandt, ¶101.) Coffman explains how the user “can instruct the PVA to dial in your PC using the phone”—i.e., use the smart phone in the user’s vehicle to establish a data connection with the user’s off-board desktop PC. (*Id.* (citing Coffman, 56:26,

57:18-20 (“you can dial in your PC”), 57:1-4 (synchronizing grocery list data from off-board desktop PC to on-board PDA via smartphone connection), 58:6-8 (“When connecting locally or through the phones....”).) A POSITA would have understood and found it obvious that using the smart phone to connect to the desktop PC in such a manner would involve “**invok[ing] a device that communicates wirelessly over a wide area network**”—i.e., the smart phone (and indirectly, the desktop PC) communicating via the public switched telephone network (PSTN) and internet. (Schmandt, ¶101.) A smart phone was a well-known device capable of communicating wirelessly over a wide area network. (*Id.* (citing EX1009, p.0032 (“smartphone”))).) More generally, Coffman describes how devices can access information “by connecting via a PSTN 1002 and internet/intranet 1003 networks.” (Coffman, 63:3-6.) The PSTN and internet are examples of a “**wide area network.**” (Schmandt, ¶101 (citing Coffman, 63:3-6; EX1009, p.0019 (“internetwork” entry describing communication over “the Internet or another WAN (wide-area network)”))).) And Coffman’s mere passing mention of “wireless network” in a list of various types of networks confirms that communication via wireless networks was well-known and conventional by at least 2001-2002. (Schmandt, ¶101 (citing Coffman, 63:14).) *See, e.g., Bayer Healthcare LLC v. Baxalta Inc.*, 989 F. 3d 964, 982 (Fed. Cir. 2021) (“[a] patent need not teach, and preferably omits, what is well known in the art”). Therefore, 1[i] is obvious. (*Id.*)

2. Claim 2: “The mobile system of claim 1, wherein executing the command or query at the vehicle comprises executing the command or query at a device built into the vehicle.”

As explained for 1[h], a POSITA would have understood and found it obvious for the PVA on-board the vehicle to execute an application call (“**the command or query**”), such as an application call for turn by turn navigation instructions. (Schmandt, ¶102.)

A POSITA would have further understood and found it obvious for the PVA device to be “**built into the vehicle.**” (Schmandt, ¶103.) Coffman’s repeated descriptions of a PVA as a “conversational car PC” never in a location other than “in a car” suggests a device built into the vehicle, in contrast to portable devices that Coffman describes, such as a “conversational PDA” and “conversational smart phone.” (*Id.* (citing Coffman, 56:21-25, 55:28).) Coffman’s further mention of a conversationally-connected “car embedded computer” further teaches a PVA built into the vehicle. (*Id.* (citing Coffman, 12:23-25 (“a car embedded computer”), 23:23-25 (same)).)

Moreover, a POSITA would have recognized that having the PVA either be “**built into the vehicle**” or somehow more generally simply “**on-board**” the vehicle would have been a non-technical design decision. (Schmandt, ¶104.) It was routine and conventional by at least 2001-2002 for car-based computers, such as interactive telematics systems, to be built into vehicles, such as in vehicle dashboards. (*Id.*

(citing EX1009, pp.0034 (“telematics” entry describing “dashboard navigation systems”), 007 (“Auto PC” entry describing device that “fits into the dashboard”); EX1006, 1:11-31 (Microsoft’s Auto PC “can fit into most automobile dashboards”)).)

3. Claim 3: “The mobile system of claim 1, wherein executing the command or query at the vehicle comprises executing the command or query at a device docked to the vehicle.”

The analysis of claim 2 addresses claim 3. A POSITA would have understood that installing Coffman’s PVA in a car’s dashboard, such as in the space normally occupied by the radio, would be “**docking**” the PVA to the vehicle. (Schmandt, ¶105 (citing EX1009, p.007 (“Auto PC”); EX1006, 1:11-31).)

Additionally, as explained for 1[h], a POSITA would have understood and found it obvious based on Coffman for a smart phone on-board the vehicle to execute an application call (“**the command or query**”), such as the application call for the phone in the user’s car to dial the user’s desktop PC. (Schmandt, ¶106.) With respect to the smart phone in the car that Coffman describes, and on-board devices more generally, it would have been obvious for the device to be “**docked**” the vehicle. (*Id.*) Whether to “dock” a device would have been a non-technical decision related to device use. (*Id.*) In-car docks for smart phones and other devices, such as holders to enable hands-free operation, were well known and commercially available by 2001-2002. (*Id.* (citing EX1009, p.0018 (“hands-free kit” describing “phone

cradle”).) Moreover, the fact that the ’845 patent recites “dock” only twice, and only in the claims without any discussion in the written description, further confirms the obviousness of “docking” a device to a vehicle. (*Id.* (citing ’845, claims 3 & 20).)

4. Claim 7: “The mobile system of claim 1, wherein executing the command or query at the vehicle comprises executing the command or query at a device that is in or on the vehicle and that is in wireless communication with the vehicle.”

With respect to “**executing the command or query at the vehicle comprises executing the command or query at a device that is in or on the vehicle,**” as explained for 1[h], a POSITA would have understood and found it obvious based on Coffman for a smart phone on-board a vehicle (“**a device that is in or on the vehicle**”) to execute an application call (“**the command or query**”), such as the application call for the phone in the user’s car to dial the user’s desktop PC. (Schmandt, ¶107.)

A POSITA would have further understood and found it obvious for the smart phone in the vehicle to be “**wireless communication with the vehicle.**” (Schmandt, ¶108.) As discussed for 1[h], Coffman describes how the PVA uses the smart phone to dial into a remotely-located desktop PC, making clear that the PVA and smart phone are communicating in some way. (*Id.* (citing Coffman, 56:26).) Coffman does not state whether the PVA and smart phone communicate via a wired or wireless connection, but it would have been obvious for the connection to be wireless. (*Id.*) Techniques and technologies for wireless communication between

devices in close proximity to each other were known and commercially available; Bluetooth is one example. (*Id.* (citing EX1009, p.009 (“Bluetooth”))). In fact, Coffman specifically identifies using Bluetooth for device communication. (Coffman, 63:14.) Using a wireless communication technology such as Bluetooth therefore would have been obvious. (*Id.*) And because as explained for claim 2, a POSITA would have understood and found it obvious for the PVA to be part of the vehicle, a POSITA would have understood and found it obvious for the smart phone in the vehicle to be **“wireless communication with the vehicle.”** (*Id.*)

5. **Claim 8: “The mobile system of claim 1, wherein the command or query includes a search query that is to be executed on-board the vehicle by searching one or more data sources that are on-board the vehicle using at least one recognized word or phrase of the natural language utterance, and wherein the search query relates to an aspect of the vehicle.”**

As discussed for claim 1, Coffman discloses a number of example conversational applications. (Coffman, e.g., 16:35-17:7, 45:2-3, 56:11-15, 57:5-9, 57:17-20, 8:24-27.) As noted, Coffman does not describe in detail a specific spoken command or query that the user could input to interact with an application, but it would have been obvious for it to include a **“search query”** as claimed—i.e., a **“search query”** that **“is to be executed on-board the vehicle by searching one or more data sources that are on-board the vehicle using at least one recognized**

word or phrase of the natural language utterance” and that **“relates to an aspect of the vehicle.”** (Schmandt, ¶109.)

For example, as discussed for 1**[b]**, with respect to vehicle navigation, Coffman describes how “[y]ou now instruct the PVA to guide you turn by turn to the [grocery] store. Your position is computed, the location of the store is fetched, locally or from a server, and an itinerary is computed to take into account the latest traffic information.” (Coffman, 57:5-7; *see also id.*, 56:18-19.) Coffman therefore teaches that executing a user’s instruction for turn by turn directions to a destination includes **“searching one or more data sources that are on-board the vehicle”** (“[T]he position of the store is fetched, locally...” (Coffman, 57:5-7)). (Schmandt, ¶110.) Coffman further teaches that the search **“relates to an aspect of the vehicle”** because it relates to providing driving directions for the vehicle, and moreover to driving directions that are based on the vehicle’s location. (*Id.*) Coffman does not detail how the “position of the store is fetched,” but a POSITA would have found it obvious to use **“at least one recognized word or phrase of the natural language utterance.”** (*Id.*) For example, it would have been obvious for a user to speak an instruction such as “driving directions to 100 Main Street” or “navigate to Market Basket,” and for the vehicle navigation application to identify an entry in a location data store by matching recognized words or phrases in the user’s utterance, such as “100,” “Main,” “Street” or “Market Basket.” (*Id.*) Searching a data store to identify

an entry by using words in a query was a well-known and conventional technique. (*Id.* (citing EX1009, pp.0021 (“keyword”), 0030 (“search string”)).) Coffman in fact expressly describes searching a data store for a matching entry using a keyword. (*Id.* (citing Coffman, 49:10-16)).)

Julia confirms the obviousness of the claimed “**search query**.” (Schmandt, ¶111.) As discussed for 1[b], Julia describes an in-car interactive navigation application that can receive a user voice command for driving directions and “take advantage of GPS information to plan a route and get location-specific information from local or remote databases displayed on the map.” (Julia, 31:18-24.) Julia explains that, “[f]or example, in response to a user voice command such as ‘I want to go to SRI on Ravenswood Avenue in Menlo Park,’ one or more GPS agents, map database agents (and/or online search agents), and multimodal map interface agents would preferably collaborate to ascertain the user’s current location and display a suggested route from there to the desired destination. and/or by speaking the suggested directions [].” (Julia, 31:24-31; *see also id.*, 31:31-35.) Julia therefore teaches a voice command that includes a “**search query**” (e.g., a search query based on the user’s utterance “I want to go to SRI on Ravenswood Avenue in Menlo Park”) and “**searching one or more data sources that are on-board the vehicle**” (e.g., “local” database) “**using at least one recognized word or phrase of the natural language utterance**” (e.g., “SRI,” “Ravenswood Avenue,” “Menlo Park”). (*Id.*)

The Petition discussed the rationale and motivation to combine Coffman and Julia in connection with 1[b], which applies here as well. (Schmandt, ¶112.) With respect to claim 8, combining Julia’s teachings regarding a vehicle navigation application with Coffman’s PVA and conversational system combination would have predictably resulted in Coffman’s navigation application searching an on-board data store using a word or phrase in the user’s voice input, such as searching a local map database using “SRI,” “Ravenswood Avenue” and “Menlo Park” in the user’s utterance “I want to go to SRI on Ravenswood Avenue in Menlo Park.” (*Id.*) The combination therefore discloses and renders obvious the claimed “**search query.**” (*Id.*)

6. Claim 10: “The mobile system of claim 8, wherein the instructions cause the one or more physical processors to:” (10[pre])

As explained below, it would have also been obvious, in view of Julia, for Coffman’s PVA to perform the claimed step (by executing computer program instructions that cause the PVA’s processor(s) to perform the step, consistent with routine and conventional operation of a generic computing device). (Schmandt, ¶113.)

- (a) “provide a natural language speech response to the natural language utterance based on one or more results from the search of the one or more data sources.” (10[a])

To start, Coffman discloses “provid[ing] a natural language speech response” as a result of processing a command/query associated with a user’s utterance. (Schmandt, ¶114.) Coffman describes “return[ing] output of executed or running processes” and presenting them to user using, for example, text-to-speech, as Coffman illustrates in Figure 6:

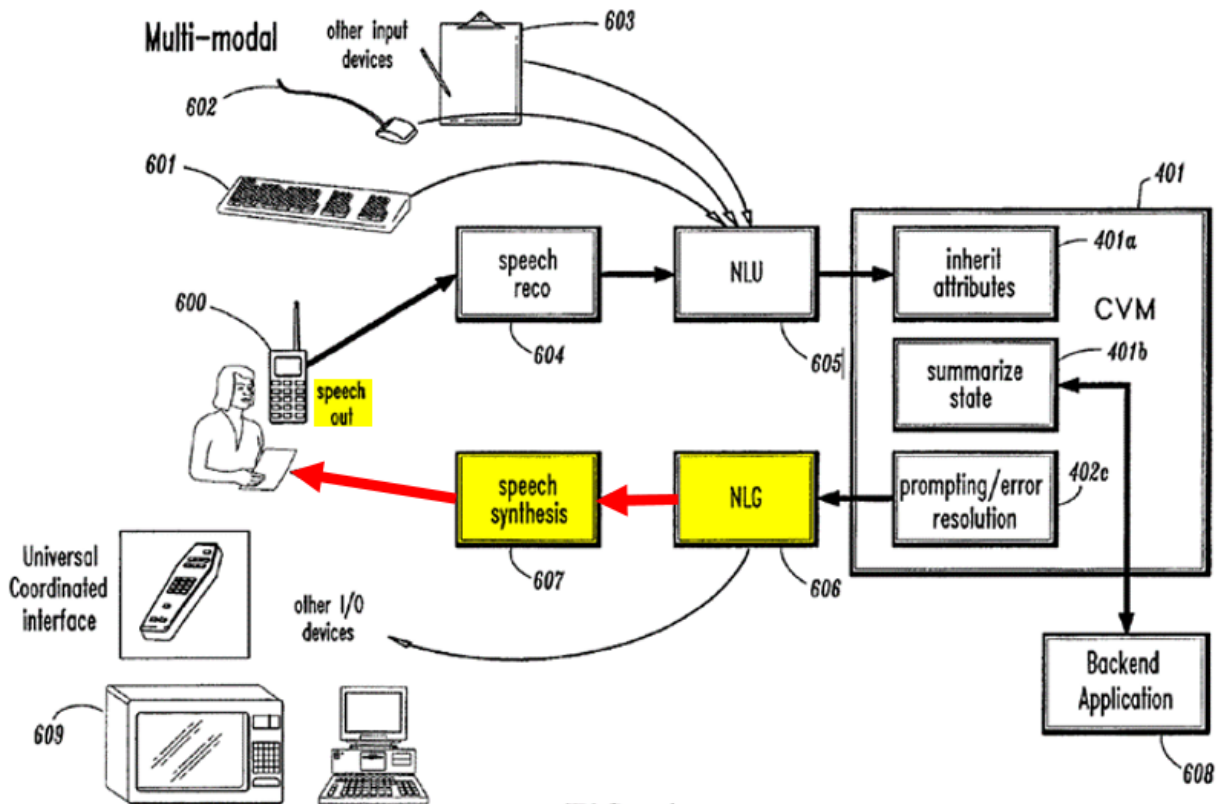


FIG. 6

(Coffman, Fig. 6 (annotations added), 40:9-19; *see also id.*, 13:8-13 & 58:9-11 (CVM converts application “output into speech...provided to the user”), 39:3-5.)

With respect to Coffman's navigation application, Coffman does not expressly describe providing driving directions in the form of natural language speech. However, as noted for claim 8, Julia, detailing a similar in-vehicle navigation application, describes providing driving directions by "display[ing] a suggested route from [the user's current location] to the desired destination. and/or by speaking the suggested directions []." (Julia, 31:24-31.) According to Julia, "[t]he system preferably guides the car along the chosen route using both the multimodal map display (e.g., to support further queries such as 'show me the restaurants along the way') and text-to-speech output 1704 for speaking the suggested driving directions aloud." (Julia, 31:31-35.) Because, as explained for claim 8, the driving directions ("**response**") are based on the results of a search of a location data store to match the user's input, they are "**based on one or more results from the search of the one or more data sources,**" as claimed. (Schmandt, ¶115.)

The rationale and motivation to combine Coffman and Julia discussed in the analysis of claim 8 applies here as well. (Schmandt, ¶116.) With respect to claim 10, the combination would have predictably resulted in Coffman's PVA "**provid[ing] a natural language speech response to the natural language utterance based on one or more results from the search of the one or more data sources.**" (*Id.*) Julia provides express motivation to incorporate, explaining that using "text-to-speech output 1704 for speaking the suggested driving directions

aloud” is “preferabl[e].” (*Id.* (quoting Julia, 31:31-35).) And more generally, a POSITA would have readily appreciated that providing spoken driving directions would advantageously allow a driver to keep their visual attention on the road, and accordingly minimize distractions and improve safety. (*Id.*)

7. Claim 12: “The mobile system of claim 1, wherein the command or query is to be executed on-board the vehicle, and wherein the command or query includes a command to control the vehicle.”

In addition to the voice-enabled navigation application discussed for claim 8, Julia also describes how “[b]y connecting a computer and its multimodal interface, the driver can control critical electronic devices such as cruise control or headlights” and that “[e]veryone in the car, including passengers, is able to access comfort devices such as air conditioning, windows, sound, and entertainment.” (Julia, 32:15-21; *see generally id.*, 32:10-26.) A command to control a vehicle’s electronic devices, such as headlights or cruise control is a command “**executed on-board the vehicle**” “**to control the vehicle.**” (Schmandt, ¶117.) Combining Julia’s teachings with Coffman’s conversational system and PVA would have predictably resulted in the user’s ability to control the vehicle by executing commands on-board the vehicle based on the user’s spoken input. (*Id.*)

The rationale and motivation to combine Coffman and Julia discussed in the analysis of 1[b] applies here as well. (Schmandt, ¶118.) Moreover, with respect to claim 12, a POSITA would have further appreciated that incorporating the ability

for the user to control the vehicle via the PVA would enhance the user's experience by further expanding the number and types of functions that the user could perform via a single voice interface. (*Id.*) Incorporating additionally functionality accessible through a single voice interface would further Coffman's recognized aims and advantages of providing a hands-free and/or eyes-free environment that coordinates the operation of multiple devices through a single interface. (*Id.* (citing Coffman, 9:12-25, 12:20-25).) Claim 12 is therefore obvious. (*Id.*)

8. Claim 15: "The mobile system of claim 1, wherein the instructions cause the one or more physical processors to:" (15[pre])

As explained below, it would have also been obvious for Coffman's PVA to perform the claimed steps (by executing computer program instructions that cause the PVA's processor(s) to perform the step, consistent with routine and conventional operation of a generic computing device). (Schmandt, ¶119.)

(a) "identify a context associated with the natural language utterance based on a current or prior dialog between the system and the user," (15[a])

The analysis of 1[e] largely addressed 15[a]. Coffman discloses "identify[ing] a context associated with the natural language utterance" by describing how a context stack is searched for an entry that matches the user's parsed utterance. (Schmandt, ¶120 (citing Coffman, 42:1-5, 41:2-5, 30:8-10, 51:12-14, 40:12-15).) Coffman further teaches that the identifying is "**based on a current or**

prior dialog between the system and the user.” (*Id.*) Coffman explains that a context “refers to the state of each discourse...which keeps track of the past history of the discourse, its current state, and the specific characteristics and full query arguments of the corresponding task....” (Coffman, 29:27-32; *see also id.*, 30:18-20 (“[E]ach context associated with a given discourse comprises the latest requests made to the corresponding process/task/thread as well as the latest output”).) Coffman further explains that as part of the context searching process, “new queries or arguments are compared by the dialog engine by going down the stack until an acceptable match is obtained and optionally confirmation is obtained from the user.” (Coffman, 42:1-5; *see also id.*, 41:2-5 (transcribed speech input “run down the context stack 405 to extract the active query and complete it.”).) Coffman therefore makes clear that the identification of a context is “**based on a current or prior dialog between the system and the user.**” (Schmandt, ¶120.)

(b) “wherein the determination of whether the command or query is to be executed on-board or off-board the vehicle is further based on the context.” (15[b])

The analysis of 1[g] largely addressed 15[b]. As explained for 1[g], Coffman teaches a distributed conversational computing system where an in-vehicle PVA acts as a master CVM that determines where—which application and device—a given command/query will be executed. (Schmandt, ¶121 (citing Coffman, 55:12-21, 55:27-30, 19:26-20:2, 56:16-57:9, 58:3-11).) And as explained for 1[e], by matching

a context, an associated application for processing the command/query is identified and activated. (*Id.* (citing Coffman, 29:18-30:3, 30:4-16, 42:1-5).) Therefore, Coffman teaches that **“the determination of whether the command or query is to be executed on-board or off-board the vehicle is further based on the context.”** (*Id.*)

9. **Claim 17: “The mobile system of claim 1, wherein the command or query includes a command that is to be executed at the vehicle, and wherein the instructions cause the one or more physical processors to:” (17[pre])**

As explained for claims 1, 8 and 10, a POSITA would have understood and found it obvious, in view of Coffman and Julia, for the user to provide a command for driving directions that would be executed by the PVA at the vehicle and result in visual and/or spoken output to the user. (Schmandt, ¶122 (citing Coffman, 57:5-7; Julia, 31:18-35).) Coffman and Julia therefore disclose and render obvious **“the command or query includes a command that is to be executed at the vehicle.”** (*Id.*)

As explained below, it would have further been obvious for Coffman’s PVA to perform the claimed steps (by executing computer program instructions that cause the PVA’s processor(s) to perform the step, consistent with routine and conventional operation of a generic computing device). (Schmandt, ¶123.)

(a) **“identify a query based on the natural language utterance;” (17[a])**

Coffman teaches that, in connection with processing the user’s command for driving directions, a query for the destination’s location is identified and carried out. (Schmandt, ¶124.) In particular, Coffman explains that the destination’s location may be “fetched, locally or from a server.” (Coffman, 57:5-7.) A POSITA would have understood that in order to “fetch” the location information, a corresponding query would be identified based on the user’s spoken input and performed. (Schmandt, ¶124.)

(b) **“invoke a device that communicates wirelessly over a wide area network to process the query such that the query is executed off-board the vehicle based on a determination that the query is to be executed off-board the vehicle; and” (17[b])**

As discussed for 1[i], Coffman describes how a PVA may “**invoke a device**” in the vehicle, such as a smart phone, that “**communicates wirelessly over a wide area network**” in connection with processing user commands and queries. (Schmandt, ¶125 (citing Coffman, 56:26, 57:18-19).) With respect to providing driving directions, a POSITA would have understood and found it obvious to invoke such a device in order for the PVA in the vehicle to be able to fetch the destination’s location “from a server” as taught by Coffman. (*Id.*) And consistent with conventional computer operation, a POSITA would have understood that “fetching” location information from the server would involve the server executing a query to

retrieve and provide the requested information. (*Id.*) Such a server-executed query, remote from the vehicle, would be “**executed off-board the vehicle.**” (*Id.*) Also consistent with conventional computer operation, a POSITA would have understood that in order for the PVA to perform operations to fetch the location information from a server, the PVA would have first determined that the server would be queried for the information—i.e., “**determine[d] that the query is to be executed off-board the vehicle.**” (*Id.*) Without making such a determination, for example, the PVA would not know to establish a communication connection or construct and transmit a properly formatted query to the server. (*Id.*)

- (c) “**receive, from a device off-board the vehicle, a response to the query, wherein the command is executed at the vehicle based on the query response.**” (17[c])

Coffman makes clear that the PVA, by “fetching” the destination’s location from a server and subsequently computing an itinerary with that information, would have “**receive[d], from a device off-board the vehicle, a response to the query.**” (Schmandt, ¶126.) And by having the PVA compute an itinerary based on that information and provide turn by turn guidance, a POSITA would have understood that the user’s command for driving directions would be “**executed at the vehicle based on the query response.**” (*Id.*)

10. Independent Claim 18

Claim 18 is the method counterpart to the system of claim 1. (Schmandt, ¶¶127-128 (side-by-side comparison).) For purposes of this Petition, there is no meaningful difference between the claims. (Schmandt, ¶128.) As explained for claim 1, Coffman’s PVA (“conversational car PC”) and conversational system is a “**computer system**”, and a POSITA would have understood and found it obvious that the recited operations would be performed by the one or more processors of the PVA executing computer program instructions. (*Id.*) Therefore, claim 18 is obvious for the same reasons as claim 1. (*Id.*)

11. Claims 19, 20, 24, 25, 27 & 29

For purposes of this Petition, claims 19, 20, 24, 25, 27 and 29 are substantially the same as claims 2, 3, 7, 8, 10 and 12, respectively. (Schmandt, ¶¶129-134.) Therefore, for the reasons discussed for claims 2, 3, 7, 8, 10 and 12, claims 19, 20, 24, 25, 27 and 29 are obvious. (*Id.*)

12. Claim 32: “A computer-implemented method of processing natural language utterances, the method being implemented by a computer system that includes one or more physical processors at a vehicle executing one or more computer program instructions which, when executed, perform the method, the method comprising:” (32[pre])

The preamble of claim 32 is identical to the preamble of independent claim 18. And similar to claim 18, claim 32 recites a method counterpart to the system of claim 1. With respect to the preamble, there is no meaningful difference between

claim 1 and claim 32. (Schmandt, ¶135.) As explained for claim 1, Coffman’s PVA and conversational system is a “**computer system**,” and a POSITA would have understood and found it obvious that the recited operations would be performed by the one or more processors of the PVA executing computer program instructions. (*Id.*) For the same reasons, a POSITA would have understood and found it obvious that the recited operations in claim 32 would be performed by the one or more processors of the PVA executing computer program instructions. (*Id.*)

- (a) “**receiving, at the one or more physical processors, a natural language utterance associated with a user;**” (32[a])
- (b) “**performing, by the one or more physical processors, speech recognition on the natural language utterance;**” (32[b])
- (c) “**parsing and interpreting, by the one or more physical processors, the speech recognized natural language utterance;**” (32[c])
- (d) “**determining, by the one or more physical processors, a domain and a context that are associated with the parsed and interpreted natural language utterance;**” (32[d])

Steps 32[a]-32[d] are identical to 18[a]-18[d]. As discussed for claim 18, these steps recite the method counterpart to 1[b]-1[e]. Therefore, they would have been obvious over the prior art for the reasons explained for 1[b]-1[e]. (Schmandt, ¶136.)

- (e) **“formulating, by the one or more physical processors, a search query based on the domain and the context wherein the search query includes at least one recognized word or phrase of the natural language utterance;” (32[e])**

32[e] differs from 1[f] in that it recites formulating a **“search query”** instead of a **“command or query,”** based on the domain and the context, and that the search query **“includes at least one recognized word or phrase of the natural language utterance.”** But as explained below, a **“search query”** as claimed would have been obvious.

As discussed for claim 1, Coffman discloses a number of example conversational applications. (Coffman, e.g., 16:35-17:7, 45:2-3, 56:11-15, 57:5-9, 57:17-20, 58:30-32, 8:24-27.) As explained, Coffman does not describe in detail a specific spoken command or query that the user could input to interact with an application, but it would have been obvious to use a spoken command or query to formulate a **“search query [that] include[d] at least one recognized word or phrase of the natural language utterance.”** (Schmandt, ¶¶138-144.)

For example, as discussed for claims 1 and 8, with respect to vehicle navigation, it would have been obvious over Coffman alone, or in view of Julia, for Coffman’s vehicle navigation application to identify information in a navigation data store by using a search query with recognized words or phrases in the user’s utterance, such as “100,” “Main,” “Street” or “Market Basket.” (Schmandt, ¶139

Petition for *Inter Partes* Review of
U.S. Patent No. 9,031,845 B2

(citing Coffman, 57:5-7, 56:18-19, 49:10-16; Julia, 30:25-45, 31:18-35; EX1009, pp.0021, 0030 (entries for “keyword” and “search string”)).)

As another example, as discussed for claim 1, Coffman describes a “conversational search capability” involving formulating a query based on the user’s spoken natural language input and then matching the query against textual information such as file name and content. (Schmandt, ¶140 (citing Coffman, 8:24-27, 51:8-15).) A POSITA would have therefore understood and found it obvious for a conversational search to be carried out by formulating a **“search query [that] include[d] at least one recognized word or phrase of the natural language utterance.”** (*Id.*) And in Coffman’s example scenario involving a distributed conversational system with a PVA as master CVM (Coffman, 56:16-58:11), a POSITA would have understood and found it obvious that a user’s spoken request to find a “grocery list” stored on the user’s desktop PC would result in the PVA instructing the desktop PC to execute a conversational search with a search query that included “grocery” or “grocery list,” and later, instructing the on-board PDA to do the same in order to have the list read back. (*Id.* (citing Coffman, 58:6-11, 57:3-4, 57:14-15).)

As another example, as discussed for claim 1, Coffman describes how “a spreadsheet, written for voice access over the desktop, can now be accessed by voice over the phone by relying on the phone capabilities.” (Coffman, 58:30-32.) In

Coffman's PVA scenario, Coffman describes how the user in the car "can dial in your PC to consult or modify some spreadsheet information." (Coffman, 57:18-20.)

A POSITA would have understood and found it obvious, consistent with ordinary use of a spreadsheet application, for a **"search query [that] include[d] at least one recognized word or phrase of the natural language utterance"** to be used in consulting or modifying spreadsheet information. (Schmandt, ¶141.) As explained for 1[b], for example, to locate a particular field or cell in a spreadsheet, it would have been obvious for the user to provide a spoken query that included words or phrases associated with the field or cell. (*Id.* (citing EX1009, p.0023 ("lookup")).) And again, in Coffman's example scenario involving a distributed conversational system with a PVA as master CVM (Coffman, 56:16-58:11), a POSITA would have understood and found it obvious for a user's spoken request to consult or modify spreadsheet information at the user's desktop PC to result in the PVA instructing the desktop PC to execute a search query. (Schmandt, ¶141 (citing Coffman, 58:6-11).)

As one more example, as discussed for 1[b], Coffman describes how a user can access their e-mail application remotely and "drive home and while driving quietly listen to your e-mail...." (Coffman, 57:17-18, 56:11-12.) As with the spreadsheet application, consistent with Coffman's teachings regarding making ordinary application functionality accessible by voice, a POSITA would have found it obvious for Coffman's system to enable a user to access their e-mail application

by voice to search for particular e-mail messages and listen to them—for example, uttering a command such as “read my e-mail messages from John Smith.” (Schmandt, ¶142.) As explained for 1[b], this is confirmed by Cooper. (6:55-58, 50:33-40, 52:1-18.) This, too, would have predictably involved a “**search query [that] include[d] at least one recognized word or phrase of the natural language utterance,**” as claimed. (Schmandt, ¶142.)

For all of these examples, the “search query” would have also been “**formulated...based on the domain and the context**” for substantially the same reasons discussed for 1[f]. (Schmandt, ¶143.) As explained, Coffman discloses formulating an application call based on matching a context and application using parsing and translation arguments provided by the associated application. (Coffman, 51:12-14, 42:1-5, 36:31-32.) With respect to 32[e], because the application call would comprise the claimed “search query,” the search query would similarly have been “**formulated...based on the domain and the context.**” (Schmandt, ¶143.)

Lastly, in addition to these examples, more generally, the use of commands comprising search queries was a well-known, routine and conventional way for a user to interact with many common software applications—such as map/navigation, spreadsheet, and email applications noted above, in addition to many others, such as word processing, web browsing, calendar and addressbook applications. (Schmandt, ¶144.) Such use was accordingly well within the background knowledge of a

POSITA. (*Id.*) Therefore, because as noted, Coffman generally teaches making ordinary application functionality accessible by voice, it would have been obvious for a user to interact with Coffman’s system by means of spoken input which would then be translated into search queries executed by Coffman’s conversationally aware applications. (*Id.*) The “**search query**” as claimed, then, would have been obvious over Coffman alone, based on the background knowledge of a POSITA. (*Id.*)

(f) **“determining, by the one or more physical processors, whether the search query is to be executed on-board or off-board a vehicle;” (32[f])**

As discussed for 1[g], Coffman describes a distributed conversational computing system having an in-vehicle PVA acting as the master CVM that determines which device should execute an application call. (Schmandt, ¶145 (citing Coffman, 19:26-20:2, 55:12-30, 56:26-57:9, 58:3-11).) As explained, in such a situation, where the architecture included devices both on-board and off-board the vehicle, a POSITA would have recognized that, by determining which device should execute the application call, the PVA would be determining whether an application call would be executed on-board or off-board the vehicle. (*Id.*)

32[f] differs from 1[g] in that it recites “**the search query**” instead of “the command or query” but the same understanding applies. (Schmandt, ¶146.) This is because, as explained for 32[e], a POSITA would have understood and found it obvious for applications both on-board and off-board the vehicle to execute a

“**search query**” as claimed. (*Id.*) For example, it would have been obvious for Coffman’s PVA to instruct an on-board navigation application or on-board name dialing application to execute a “**search query.**” (*Id.*) Similarly, it would have been obvious for the PVA to instruct an off-board device, such as a user’s desktop PC, to carry out a conversational search, or to consult or modify spreadsheet information, or to locate particular e-mail messages by executing a “**search query.**” (*Id.*) Therefore, in a distributed conversational computing system having an in-vehicle PVA acting as the master CVM as taught by Coffman, a POSITA would have recognized that, by determining which device should execute the application call involving a search query, the PVA would be determining whether the search query would be executed on-board or off-board the vehicle (“**determining, by the one or more physical processors, whether the search query is to be executed on-board or off-board a vehicle**”). (*Id.*)

- (g) “**executing, by the one or more physical processors, the search query at the vehicle in response to a determination that the search query is to be executed on-board the vehicle; and**” (32[g])

As explained for 32[e], a POSITA would have understood and found it obvious for a device on-board the vehicle, such as the PVA, to execute the search query, such as a search query for the PVA’s navigation application to execute. (Schmandt, ¶147.) As explained for 32[f], it also would have been understood and obvious for the search query to be executed in response to the PVA determining

where the search query should be executed (“**in response to a determination that the search query is to be executed on-board the vehicle**”). (*Id.*)

- (h) “**invoking, by the one or more physical processors, a device that communicates wirelessly over a wide area network to process the search query such that the search query is executed off-board the vehicle in response to a determination that the search query is to be executed off-board the vehicle.**” (32[h])

32[h] is obvious for substantially the same reasons as 1[i]. (Schmandt, ¶148.)

As explained for 32[e], it would have been obvious that an application call to an off-board device would involve a “**search query**” as claimed. (*Id.*) Therefore, a POSITA would have understood that the PVA determining where the search query should be executed would predictably result in the search query being “**executed off-board the vehicle in response to a determination that the search query is to be executed off-board the vehicle.**” (*Id.*)

And as discussed for 1[i], Coffman describes how a PVA invokes a user’s smart phone (“**a device that communicates wirelessly over a wide area network**”) to communicate with the user’s desktop PC off-board the vehicle. (Coffman, 56:26, 57:18-20.) A POSITA would have therefore understood and found it obvious that, by using the smart phone to communicate with the off-board desktop PC, Coffman’s PVA would be “**invoking, by the one or more physical processors, a device that communicates wirelessly over a wide area network to process the search query.**” (Schmandt, ¶149.)

13. Claims 34 & 36

For purposes of this Petition, claims 34 and 36 are substantially the same as claims 8 and 10, respectively. (Schmandt, ¶¶150-151.) Therefore, for the reasons discussed for claims 8 and 10, claims 34 and 36 are obvious. (*Id.*)

VII. CONCLUSION

Petitioner respectfully requests IPR institution.

Dated: April 5, 2024

Respectfully submitted,

COOLEY LLP
ATTN: Patent Group
1299 Pennsylvania Avenue NW
Suite 700
Washington, DC 20004
Tel: (650) 843-5001
Fax: (650) 849-7400

By: / Heidi L. Keefe /
Heidi L. Keefe
Reg. No. 40,673
Counsel for Petitioner

CERTIFICATE OF COMPLIANCE WITH WORD COUNT

Pursuant to 37 C.F.R. § 42.24(d), I certify that this petition complies with the type-volume limits of 37 C.F.R. § 42.24(a)(1)(i) because it contains 13,983 words, according to the word-processing system used to prepare this petition, excluding the parts of this petition that are exempted by 37 C.F.R. § 42.24(a) (including the table of contents, a table of authorities, mandatory notices, a certificate of service or this certificate word count, appendix of exhibits, and claim listings).

DATED: April 5, 2024

COOLEY LLP
ATTN: Patent Docketing
1299 Pennsylvania Avenue NW
Suite 700
Washington, D.C. 20004
Tel: (650) 843-5001
Fax: (650) 849-7400

/ Heidi L. Keefe /
Heidi L. Keefe
Reg. No. 40,673

CERTIFICATE OF SERVICE

I hereby certify, pursuant to 37 C.F.R. Sections 42.6 and 42.105, that a complete copy of the attached **PETITION FOR INTER PARTES REVIEW OF U.S. PATENT NO. 9,031,845 B2**, including all exhibits (**Nos. 1001-1006, 1008-1009, 1018-1020**) and related documents, are being served via Federal Express on the 5th day of April, 2024, the same day as the filing of the above-identified document in the United States Patent and Trademark Office/Patent Trial and Appeal Board, upon Patent Owner by serving the correspondence address of record with the USPTO as follows:

174899 - The Law Offices of David A. Gerasimow, P.C.
211 W. Wacker Dr.
Ste. 1717
Chicago, IL 60606
UNITED STATES

And, via Federal Express upon counsel of record for Patent Owner in the litigation pending before the U.S. District Court for the District of Delaware entitled *Dialect, LLC v. Google LLC*, Case No. 1:23-cv-00378 (D. Del.) as follows:

Garland Stephens
BLUE PEAK LAW GROUP LLP
3139 W Holcombe Blvd, PMB 8160
Houston, TX 77025
Telephone: 281-972-3036
garland@bluepeak.law

DATED: April 5, 2024

/ Heidi L. Keefe /
Heidi L. Keefe
Reg. No. 40,673
COOLEY LLP
1299 Pennsylvania Ave. NW
Suite 700
Washington, D.C. 20004
Tel: (650) 843-5001