

Preliminary Ineligibility Contentions for U.S. Patent 10,755,699

Pursuant to Chief Judge Gilstrap's Standing Order Regarding Subject Matter Eligibility Contentions, the following charts contain Samsung's Preliminary Ineligibility Contentions demonstrating that claims 1-22 of U.S. Patent No. 10,755,699 ("'699 patent") are patent-ineligible under 35 U.S.C. § 101. Because the Court has yet to issue a claim construction in this case, and to the extent Plaintiff contends that there are factual disputes that relate to eligibility of the asserted claims of the '699 patent, fact and expert discovery is ongoing, these contentions are preliminary only and Samsung reserves the right to supplement or modify these contentions in accordance with the agreed patent-disclosure procedures and the Docket Control Order in this case. Additionally, and in further consideration of the preliminary stages of the case, Samsung notes that the pinpoint citations referenced in this chart are not exhaustive, and Samsung reserves the right to rely on additional citations within the asserted patent any cited reference. Furthermore, citations to any figure, table, or chart are meant to encompass the language describing the figure, table, or chart, and vice versa. To the extent applicable, Samsung incorporates by reference its citations in its eligibility contentions for the other asserted patents, which are related to the '699 patent.

To the extent applicable, Samsung further incorporates by reference its citation in the § 102/103 contentions, which illustrate that the claimed elements were well-understood, routine, and conventional.

Further, these charts may incorporate Plaintiff's apparent interpretation of the breakdown of elements within the asserted claims, as reflected in Plaintiff's infringement contentions to date. Samsung does not concede that such interpretation is correct, and reserves its right to supplement these contentions accordingly.

Samsung reserves its right to supplement these contentions based on further discovery, including any supplemental infringement contentions or to any interrogatory response provided by Plaintiff purporting to rebut these ineligibility contentions.

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

Pursuant to Judge Gilstrap’s Standing Order Regarding Subject Matter Eligibility Contentions, Samsung provides the following chart identifying each exception to eligibility (e.g., abstract idea, law of nature, and natural phenomenon) to which each Challenged Claim is directed and the factual and legal basis therefor, as well as identification of whether one or more of the Challenged Claims are representative of any other Challenged Claims. Samsung incorporates by reference the prior art identified in the ’699 patent, the prosecution history of the ’699 patent, any related litigations, and any related PTAB proceedings. Samsung further incorporates by reference any prior art identified or to be identified in its invalidity contentions and/or expert reports that will be provided according to the schedule provided by the Court’s Docket Control Order (Dkt. No. 42).

| Challenged Claims | Exception to Eligibility |
|--|--|
| <p style="text-align: center;">Claims 1-22</p> | <p>Each of the independent claims of the ’699 patent are directed to the abstract idea of interpreting and responding to words based on the manner in which the words were spoken. This is not a technological advancement in the art, but, as the patent discloses, a basic feature of all human communication. ’699 patent, 1:51-53, 1:35-51.</p> <p>Accordingly, the patent is subject to the “abstract idea” exception to patent eligibility under the Patent Act as construed by the Supreme Court in <i>Alice</i>.</p> <p><u>Representativeness</u></p> <p>As an initial matter, for the purpose of the <i>Alice</i> inquiry, Claim 1 is representative of all asserted claims of the ’699 patent. VBA asserts independent claims 1 and 12 and dependent claims 2-11 and 13-22.</p> <p>Starting with the independent claims, the claims are identical except for that claim 1 claims a computer-implemented method and claim 13 claims a system. This is displayed in the following table:</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility | | |
|-------------------|---|---------------------|----------------------|
| | Feature | Claim 1 (Method) | Claim 13 (System) |
| | Receives an utterance | ✓ | ✓ |
| | Recognizes word(s) from the utterance | ✓ | ✓ |
| | Identifies a context for the utterance | ✓ | ✓ |
| | Determines an interpretation based on the context | ✓ | ✓ |
| | Accumulates short-term shared knowledge | ✓ | ✓ |
| | Accumulates long-term shared knowledge | ✓ | ✓ |
| | Identifies a manner in which the utterance was spoken | ✓ | ✓ |
| | Generates a response | ✓ | ✓ |
| | <p>The dependent claims do not alter the analysis as the dependent claims are substantially similar, if not substantially identical, and linked to the same abstract idea. <i>See Content Extraction & Transmission LLC v. Wells Fargo Bank, Nat. Ass’n</i>, 776 F.3d 1343, 1348 (Fed. Cir. 2014).</p> <p>For example, Claims 2 and 13 claim variants of Independent Claims 1 and 12, where identified manner includes an indication of “tone, pace, timing, inflection, word use, and/or jargon.”</p> <p>Claims 3, 9-11, 13, and 14-22 limit how the response of Claims 1 and 12 is generated and adapted. Claims 3 and 14 similarly limit the response by requiring a variation of “tone, pace,</p> | | |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility |
|-------------------|---|
| | <p>timing, inflection, word use, and/or jargon of the voice response.” Claims 9 and 20 specify that the response is a voice response and require varying word use and/or jargon. Claims 10 and 21 require the response to model a conversation, which is achieved by adapting the response to have a personality by varying word use. And Claims 11 and 22 require the response is “sensitive to context, what the user already knows about a topic, short-term knowledge and long-term knowledge of user preferences, and words uttered by the user in one or more prior natural language utterances.”</p> <p>Claims 4 and 15 add the steps of obtaining contextual signifiers and/or grammatical rules from the utterance and using those signifiers and rules to generate sentences as part of a response set to cooperate with the user.</p> <p>Claims 5 and 16 further narrow Claims 1 and 12 in that they provide for generating a user profile based on the long-term knowledge, which is then used to help determine the context of the utterance.</p> <p>Claims 6 and 17 require that the response in Claims 1 and 12 is adapted based on a response format associated with the manner.</p> <p>Claims 7 and 18 specify that the utterance is an input to a speech recognition engine and the recognized words/phrases are obtained as an output of the speech recognition engine.</p> <p>Claims 8 and 19 add the inconsequential step of providing the generated response to the user.</p> <p>As detailed above, all of the asserted dependent claims merely narrow the scope of the independent claims to use specific modalities or decision processes for obtaining the same abstract idea claimed in the independent claims. Because these limitations do not provide any additional inventive concept, their eligibility under 101 rises and falls with the independent claims.</p> |

| Challenged Claims | Exception to Eligibility |
|-------------------|---|
| | <p><u>The ’699 Patent Claims an Abstract Idea</u></p> <p>In <i>Beteiro, LLC v. DraftKings Inc.</i>, the Federal Circuit recently identified certain “well-settled indicators of abstractness.” 104 F. 4th 1350, 1355 (Fed. Cir. 2024). These are: (1) claims that broadly recite generic steps; (2) claims drafted using largely result-focused functional language, and (3) claims that can be analogized to “longstanding ‘real-world’ activities. <i>Id.</i> While each of these indicators are sufficient to demonstrate abstractness, the claims of the ’765 patent exhibit all of the features identified by that court.</p> <p><i>First</i>, the claims are broad and cite generic steps. In considering the breadth and genericness of the claims, the focus must be on the actual claims. <i>See ChargePoint, Inc. v. SemaConnect, Inc.</i>, 920 F.3d 759, 769 (Fed. Cir. 2019) (“[T]he § 101 inquiry must focus on the language of the Asserted Claims themselves” (quoting <i>Synopsys, Inc. v. Mentor Graphics Corp.</i>, 839 F.3d 1138, 1149 (Fed. Cir. 2016))); <i>Alice</i>, 573 U.S. at 219, 134 S.Ct. 2347 (“On their face, the claims before us are drawn to the concept of intermediated settlement” (emphasis added)). When evaluating the claim language, the court must evaluate “the ‘focus of the claimed advance over the prior art’ to determine if the claim’s ‘character as a whole’ is directed to excluded subject matter.” <i>Intell. Ventures I LLC v. Erie Indem. Co.</i>, 850 F.3d 1315, 1325 (Fed. Cir. 2017) (quoting <i>Affinity Labs of Tex., LLC v. DIRECTV, LLC</i>, 838 F.3d 1253, 1257 (Fed. Cir. 2016)).</p> <p>While claim 1 of the ’699 patent includes extraneous verbiage designed to provide a perception of technological advance and complexity of the alleged invention in an attempt to remove it from the realm of ineligible subject matter, it boils down to a claim over the following six-step process: (1) receiving an utterance, (2) using recognized words to determine a context for the utterance, (3) determining an interpretation of the utterance in view of the context, (4) acquiring and storing information about the current and former conversations, (5) identifying a manner in which the utterance was spoken in view of the current and former conversation, and (6)</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants' Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility |
|--------------------------|--|
| | responding based on the interpretation and manner of the utterance. This mapping of the patent elements to the claimed concept is exemplified below: |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants' Invalidation Contentions – Appendix G-3

| | '699 patent | Claimed Concept |
|--|--|--|
| | [1pre]: A computer-implemented method of generating natural language system responses adapted based on a user's manner of speaking, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising: | A method for speech recognition, comprising: |
| | [1A] receiving, by the computer system, a user input comprising a natural language utterance; | Step 1: Receiving an utterance; |
| | [1B] recognizing, by the computer system, one or more words or phrases from the natural language utterance; | Step 2: Recognizing words; and |
| | [1C] identifying, by the computer system, a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance; | Step 2: Using those words to identify a context for the utterance; |
| | [1D] determining, by the computer system, an interpretation of the natural language utterance based on the identified context; | Step 3: Determining an interpretation of the utterance in view of the context; |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility | |
|-------------------|--|--|
| | [1E] accumulating, by the computer system, short-term knowledge based on one or more natural language utterances received during a predetermined time period, | Step 4: Acquiring and storing information about utterances received during a specific time period; where |
| | [1E.1] wherein the one or more natural language utterances received during the predetermined time period are related to a single conversation between a user and the computer system; | Step 4: The time period captures a single conversation; |
| | [1F] accumulating, by the computer system, long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period; | Step 4: Acquiring and storing information about former conversations; |
| | [1G] identifying, by the computer system, a manner in which the natural language utterance was spoken based on the short-term knowledge and the long-term knowledge; and | Step 5: Identifying a manner in which the utterance was spoken in view of the current and former conversation; |
| | [1H] generating, by the computer system, a response to the natural language utterance based on the interpretation and the identified manner in which the natural language utterance was spoken. | Step 6: Responding based on the interpretation and manner of the utterance. |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility |
|-------------------|--|
| | <p>These steps—receiving a message, detecting information, and generating and transmitting a response—are exactly the type of steps that the Federal Circuit has “frequently” held to be abstract. <i>See, e.g., AI Visualize, Inc. v. Nuance Commc’ns, Inc.</i>, 97 F.4th 1371, 1378 (Fed. Cir. 2024) (“We have explained that the steps of obtaining, manipulating, and displaying data, particularly when claimed at a high level of generality, are abstract concepts.”); <i>Elec. Power Grp. LLC v. Alstom S.A.</i>, 830 F.3d 1350, 1354 (Fed Cir. 2016) (finding “a process of gathering and analyzing information of a specified content, then displaying the results” abstract); <i>Intell. Ventures I LLC v. Capital One Bank (USA)</i>, 792 F.3d 1363, 1369 (Fed. Cir. 2015) (finding “tailoring content based on the [user’s] location” abstract); <i>buySAFE, Inc. v. Google, Inc.</i>, 765 F.3d 1350, 1355 (Fed. Cir. 2014) (claiming “a computer receiv[ing] and send[ing] information over a network – with no further specification – is not even arguably inventive”).</p> <p><i>Second</i>, the claims are drafted using largely result-focused functional language that “contain[s] no specificity about how the purported invention achieves those results.” <i>Beteiro, LLC v. DraftKings Inc.</i>, 104 F.4th 1350, 1356 (Fed. Cir. 2024). As the Federal Circuit has observed “Claims of this nature are almost always found to be ineligible for patenting under Section 101.” <i>Id.</i> (citing <i>Elec. Power Grp.</i>, 830 F.3d at 1356 (“[T]he essentially result-focused, functional character of claim language has been a frequent feature of claims held ineligible under § 101, especially in the area of using generic computer and network technology to carry out economic transactions.”); <i>Two-Way Media Ltd. v. Comcast Cable Commc’ns, LLC</i>, 874 F.3d 1329, 1337 (Fed. Cir. 2017) (“The claim requires the functional results of ‘converting,’ ‘routing,’ ‘controlling,’ ‘monitoring,’ and ‘accumulating records,’ but does not sufficiently describe how to achieve these results in a non-abstract way.”)).</p> <p>Namely, as illustrated with Claim 1, the claims recite that utterances are “received”, without disclosing any technical insight into how they are received; that words are “recognized,” without any insight into how the words are recognized; that a context is “identified” based on the recognized words, without any insight on how the words are associated with a context; that an interpretation is “determined” from the identified context, without and insight on how the determination is made; that knowledge is “accumulate[ed],” without any technical insight into</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility | |
|-------------------|---|------------------------------|
| | <p>how that knowledge is accumulated; that manner is “identified” from the short- and long-term knowledge, without any insight into how the identification is made; and that a response is “generated,” in a matter tailored to the interpretation and manner, without any insight into how that response is generated. At no point do the inventors disclose “how” these steps are accomplished.</p> <p><i>Third</i>, the claims are directly analogous to “longstanding ‘real-world’ (‘brick-and-mortar’) activities.” <i>Beteiro, LLC v. DraftKings Inc.</i>, 104 F.4th at 1356; <i>see also PersonalWeb Techs. LLC v. Google LLC</i>, 8 F.4th 1310, 1316 (Fed. Cir. 2021) (claims that “can be performed in the human mind” or “using a pencil and paper” are “a telltale sign of abstraction”); <i>Intell. Ventures I LLC v. Symantec Corp.</i>, 838 F.3d 1307, 1314 (Fed. Cir. 2016) (a claimed method’s similarity to “fundamental ... practices long prevalent” is yet another clue that the claims may be abstract and unpatentable). For example, since the beginning of language, humans and their brains have been parsing both short-term knowledge and long-term knowledge to facilitate fluent conversations. Short-term knowledge, such as the narrow topic area being discussed at the present time, does and always has driven a conversation forward. Likewise, long-term knowledge, such as the identity of the person speaking who was first met years ago, is used to influence how we respond and communicate. As one specific example in the law firm setting demonstrates, the claims read on a junior associate’s response to a senior partner’s query of “Is the Claim Chart you’ve now been working on for 3 weeks done yet?”</p> | |
| | ’699 patent | Everyday Conversation |
| | <p>[1pre]: A computer-implemented method of generating natural language system responses adapted based on a user's manner of speaking, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which,</p> | |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility | |
|-------------------|--|--|
| | when executed, perform the method, the method comprising: | |
| | [1A] receiving, by the computer system, a user input comprising a natural language utterance; | Individual receives an utterance: “Is the Claim Chart you’ve been working on for 3 weeks done yet?” |
| | [1B] recognizing, by the computer system, one or more words or phrases from the natural language utterance; | Individual recognizes words/phrases, including: - “Claim Chart” - “three weeks” - “done.” |
| | [1C] identifying, by the computer system, a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance; | Identifying that the context for the utterance is associated with the work that the Individual had been doing for the speaker on the ’699 patent in the <i>VB Assets</i> case. |
| | [1D] determining, by the computer system, an interpretation of the natural language utterance based on the identified context; | Determining that the speaker was asking whether the Individual’s work on the preliminary ineligibility claim chart for the ’699 patent had been completed. |
| | [1E] accumulating, by the computer system, short-term knowledge based on one or more natural language utterances received during a predetermined time period, | Individual acquires short-term knowledge about the manner in which the request was spoken, including: - Tone - Pace |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility | |
|-------------------|---|---|
| | | <ul style="list-style-type: none"> - Timing - Inflection - Word use - Jargon. |
| | <p>[1E.1] wherein the one or more natural language utterances received during the predetermined time period are related to a single conversation between a user and the computer system;</p> | |
| | <p>[1F] accumulating, by the computer system, long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period;</p> | <p>Individual has long-term knowledge about the speaker, including:</p> <ul style="list-style-type: none"> -The identity of the speaker -The due date of the Claim Chart that the receiver had been working on for the speaker. |
| | <p>[1G] identifying, by the computer system, a manner in which the natural language utterance was spoken based on the short-term knowledge and the long-term knowledge; and</p> | <p>Identifying that the manner in which the speaker made the request was irritated based on the harsh tone, the inclusion of the words/phrases “three weeks” and “yet,” and the knowledge that the Claim Chart was due one week ago.</p> |
| | <p>[1H] generating, by the computer system, a response to the natural language utterance based on the interpretation and the identified</p> | <p>Responding: “No, not yet. I apologize for missing the deadline and will revert it to you by the end of the day.”</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Exception to Eligibility | |
|-------------------|---|--|
| | <p>manner in which the natural language utterance was spoken.</p> | <p>That the patent claims performing this process using electronic devices and speech recognition software or hardware does not change the ultimate conclusion that the patent is directed to the abstract idea of improving natural language processing using conversational principles. Indeed, the Federal Circuit has repeatedly held that a patent is directed to an abstract idea when the claimed invention simply automates via computer a process that has been and can be effectuated by humans. <i>See CyberSource Corp., v. Retail Decisions, Inc.</i>, 654 F.3d 1366, 1373 (Fed. Cir. 2011); <i>see also Voter Verified, Inc. v. Election Sys. & Software LLC</i>, 887 F.3d 1376, 1385 (Fed. Cir. 2018) (holding patent claims directed to auto-verification of a ballot to be abstract because “[h]umans have performed [the act of voting] . . . for hundreds of years.”); <i>Repifi Vendor Logistics, Inc. v. IntelliCentrics, Inc.</i>, 2022 WL 794981, *2-3 (Fed. Cir. 2022) (claims were “directed to the abstract idea of credentialing visitors and checking them in and out”—a “long-standing human process” (quotation marks omitted)); <i>CardioNet, LLC v. InfoBionic, Inc.</i>, 816 F. App’x 471, 476–77 (Fed. Cir. 2020) (measuring heart data claims abstract because process “could otherwise be performed by a human”); <i>Perry St. Software, Inc. v. Jedi Techs., Inc.</i>, 548 F. Supp. 3d 418, 432 (S.D.N.Y. 2021) (“Jedi seeks to apply these age-old processes to the context of internet chatrooms. But facilitating human interactions or relationships – which could occur offline and without automation – is not a patent-eligible idea. It is, as Judge Sleet found in connection with Jedi's earlier patents, nothing more than the mechanization of traditional processes designed to satisfy humanity's desire for companionship.”)</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

Pursuant to Judge Gilstrap’s Standing Order Regarding Subject Matter Eligibility Contentions, Samsung provides the following chart identifying a description of the industry, at the relevant time, in which the Challenged Claims are alleged to be well understood, routine, and conventional, and the factual and legal basis therefor; and a description of how each element of each Challenged Claim, both individually and in combination with the other elements of that claim, was well understood, routine, and conventional, in the relevant industry at the relevant time, and the legal and factual basis therefor.¹

The claims of the 699 patent amount to nothing more than the practice of an abstract idea using conventional (even as of 2006) electronics. Such claims are not eligible for patent under current Section 101 jurisprudence. *See, e.g., Alice, 573 U.S. at 223, 134 S.Ct. 2347* (“[T]he mere recitation of a generic computer cannot transform a patent-ineligible abstract idea into a patent-eligible invention.”); *BSG Tech.*, 899 F.3d at 1290-91 (“If a claim’s only ‘inventive concept’ is the application of an abstract idea using conventional and well-understood techniques, the claim has not been transformed into a patent-eligible application of an abstract idea.”). Accordingly, the district court rightly dismissed the complaints.

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| <p>[1P] A computer-implemented method of generating natural language system responses adapted based on a user's manner of speaking, the method being implemented by a computer system that includes one or more physical processors executing one or more computer program instructions which, when executed, perform the method, the method comprising:</p> | <p>The claimed computer-implemented method of generating natural language system responses adapted based on a user's manner of speaking, implemented by a computer system with one or more physical processors executing program instructions, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically adapt system responses based on a user's manner of speaking beyond high-level conceptual statements. This lack of technical detail indicates that these</p> |

¹ Samsung notes that under Supreme Court and Federal Circuit precedent, the Court need not, and should not, consider whether each and every element of each Challenged Claim was well understood, routine, and conventional. Rather, the only elements which should be considered under *Alice* step 2 are the elements that fall outside the scope of the abstract idea. *See Alice Corp. v. CLS Bank Int’l*, 573 U.S. 208, 221-224 (2014); *BSG Tech LLC v. BuySeasons, Inc.*, 899 F.3d 1281, 1290 (Fed. Cir. 2018) (“It has been clear since *Alice* that a claimed invention’s use of the ineligible concept to which it is directed cannot supply the inventive concept that renders the invention ‘significantly more’ than that ineligible concept.”).

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>The adaptation of system responses based on detected aspects of a user’s speech patterns, including prosodic features such as pitch, timing, energy, and speaking style, was a well-established technique. For example, Hirschberg’s <i>Communication and Prosody: Functional Aspects of Prosody</i> (1999) discusses how users vary prosody naturally in communication and how systems can exploit prosodic cues such as pitch range and rate to detect user engagement or emotional state. Additionally, Wahlster’s <i>SmartKom: Foundations of Multimodal Dialogue Systems</i> (2006) describes multimodal dialogue systems that analyze user prosody to adjust dialogue strategies dynamically based on detected emotions or speaking behaviors.</p> <p>The intrinsic record is consistent, it describes:</p> <p>Freeman and Baldwin (2006) disclose how cooperative conversation concepts were first described by Paul Grice in 1975 and how human-to-computer speech should ideally “follow the same model of human-to-human conversation” described in <i>Logic and Conversation</i>. Freeman and Baldwin explain that a system’s ability to modify responses based on user input falls under the general category of an adaptive response:</p> <p>Adaptive Responses</p> <p>"As in a human conversation, the response of the system should be based on how confident the system is that its conclusion is correct. If it is certain, it should just execute the request. If it is less than certain but the</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>range of choice is few or none, it should execute. If it is less than certain and the choices are many, it should ask."</p> <p>Tom Freeman and Larry Baldwin, Enhancing the VUE™ (Voice-User-Experience) Through Conversational Speech, SAE Technical Paper Series (2006).</p> <p>This framework, described in 2006, demonstrates that adapting responses based on user input—including manner of speaking—was well known and widely understood in the field.</p> <p>Moreover, under oath, Mr. Freeman explicitly confirmed that the principles set forth in Freeman and Baldwin (2006), including the adaptive response mechanism, correspond to the limitations described in the '681 patent—the parent to the '699 patent:</p> <p style="padding-left: 40px;">Q. And this marketing paper where you describe these principles that Mr. Grice had laid out, this is essentially the '681 patent in ordinary language, right?</p> <p style="padding-left: 40px;">A. So this is '681 but '681 takes this to a machine, as opposed to human to human.</p> <p>(Trial Tr. 289:20-24, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>To the extent that the claimed method simply adapts system responses based on a user's speaking manner, it falls squarely within this well-established adaptive response framework and does not introduce any novel technological improvement.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>Further, Paul Grice in fact emphasized how the cooperative principal required consideration of prosody to understand meaning. He specifically explicated on the topic in certain “Further Notes” he provided on “Logic and Conversation.” There, he observed that there was more to consider in utterances than just the utterances themselves, and that “[w]e should consider also examples of elements in or aspects of utterances which, not being words, are candidates for conventional meaning.” For instance, he described, utterances can be made to stress certain syllables or words, or with certain tones, such as those that connote irony.</p> <p>Further, the specification of the ’699 patent does not disclose any technical details or implementation specifics regarding how the system actually adapts responses based on a user’s manner of speaking. Instead, it provides broad, functional language:</p> <p style="padding-left: 40px;">“According to another aspect of the invention, the intelligent responses may adapt to a user's manner of speaking by using contextual signifiers and grammatical rules to generate one or more sentences that may cooperate with the user. By taking advantage of shared knowledge about how a user utters a request, the responses may be modeled using similar techniques used to recognize requests. The intelligent responses may rate possible responses statistically and/or randomize responses, which creates an opportunity to build an exchange with natural variation and conversational feel. This provides advantages over existing voice user interfaces where input and output is incongruous, as the input is ‘conversational’ and the output is ‘computerese.’” (’699 patent at 6:31-43).</p> <p>This passage merely describes desired outcomes—that system responses “may adapt” to a user’s speech patterns—but provides no specific algorithm or method</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|---|
| | <p>for achieving this adaptation. It does not explain how contextual signifiers are identified, how grammatical rules are applied dynamically, or how statistical weighting is implemented in generating adaptive responses. Instead, it broadly asserts that responses may be modeled similarly to request recognition, without detailing any underlying technological mechanism.</p> <p>Thus, the claimed method of generating natural language system responses that adapt based on a user's manner of speaking was well understood, routine, and conventional in the relevant industry at the time of filing, and the specification provides no meaningful details on how this adaptation is technically accomplished.</p> <p>Finally, the '699 patent specification provides no meaningful disclosure as to the hardware or software used to implement the claims. To the extent any hardware is disclosed, it is generic, consisting of devices such as a microphone to receive the utterance or a speaker to deliver the generated response. '699 patent at 7:36-38, 28:18-21. Such devices were standard technology at the time of the invention, and thus, the claims of the '699 patent do not recite any improvement in computer technology.</p> |
| <p>[1A] receiving, by the computer system, a user input comprising a natural language utterance;</p> | <p>The claimed method of receiving, by the computer system, a user input comprising a natural language utterance was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically receive a natural language utterance as user input beyond generally describing the concept. This lack of technical detail indicates that these processes</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Receiving user input in the form of natural language utterances was a fundamental practice in dialogue systems and speech-based user interfaces. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes systems that process speech-based natural language input from users to initiate and manage conversational interactions.</p> <p>For example, the intrinsic record describes how:</p> <p style="padding-left: 40px;">“The system may receive an input 105 from a user, where in one implementation, input 105 may be an utterance received by an input device (e.g., a microphone), where the utterance may include one or more requests. Input 105 may also be a multi-modal input, where at least part of the multi-modal input is an utterance.” (’699 patent at 7:36-41).</p> <p>Additionally, the specification confirms that:</p> <p style="padding-left: 40px;">“The utterance component of input 105 may be processed by a speech recognition engine 110 (which may alternatively be referred to herein as Automatic Speech Recognizer 110, or as shown in FIG. 1, ASR 110) to generate one or more preliminary interpretations of the utterance. The speech recognition engine 110 may process the utterance using any suitable technique known in the art.” (’699 patent at 7:50-56).</p> <p>The ’699 patent at 2:2-8 further acknowledges that existing speech interfaces had long processed user utterances as input:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>“Existing speech interfaces force users to dumb down their requests to match simple sets of instructions in simple languages in order to communicate requests in ways that systems can understand. Using existing speech interfaces, there is virtually no option for dialogue between the user and the system to satisfy mutual goals.”</p> <p>Further, there is no technical disclosure or explanation in the specification about how a system specifically receives and processes user utterances beyond the broad assertion that it functions similarly to existing speech recognition technologies. The patent does not provide a specific algorithm, process, or computational method for differentiating between various types of utterances or handling ambiguous input. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of receiving a user input comprising a natural language utterance was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| [1B] | <p>The claimed method of recognizing, by the computer system, one or more words or phrases from the natural language utterance was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically recognize words or phrases beyond a general reference to speech recognition processes. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>The recognition of words and phrases from natural language utterances through automatic speech recognition (ASR) systems was a well-established and fundamental technique. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes numerous spoken dialogue systems that process user speech input by recognizing individual words and phrases to drive dialogue management.</p> <p>Consistently, the intrinsic record describes how:</p> <p style="padding-left: 40px;">“The utterance component of input 105 may be processed by a speech recognition engine 110 (which may alternatively be referred to herein as Automatic Speech Recognizer 110, or as shown in FIG. 1, ASR 110) to generate one or more preliminary interpretations of the utterance. The speech recognition engine 110 may process the utterance using any suitable technique known in the art.” (’699 patent at 7:50-56).</p> <p>Additionally, the specification confirms that:</p> <p>“speech recognition engine 110 may interpret the utterance using techniques of phonetic dictation to recognize a stream of phonemes, as described in U.S. patent application Ser. No. 11/513,269, entitled ‘Dynamic Speech Sharpening’ filed Aug. 31, 2006, which issued as U.S. Pat. No. 7,634,409 on Dec. 15, 2009.” (’699 patent at 7:57-62).</p> <p>Moreover, the inventors of the ’699 patent have admitted under oath that the ASR technology they employed was a “commodity” product at the time.</p> <p>Further, there is no technical disclosure or explanation in the specification about how a system specifically recognizes words or phrases beyond the broad assertion that existing speech recognition techniques are used. The patent does not provide a specific algorithm, process, or computational method for improving recognition</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| | <p>accuracy, handling ambiguous input, or distinguishing between homophones. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of recognizing words or phrases from a natural language utterance was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[1C] identifying, by the computer system, a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance;</p> | <p>The claimed method of identifying, by the computer system, a context for the natural language utterance based on the one or more words or phrases recognized from the utterance was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically determine context based on recognized words or phrases beyond high-level conceptual language. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Determining conversational or situational context based on recognized linguistic content from user utterances was a well-established approach in dialogue systems and natural language understanding. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes dialogue systems that infer dialogue context and conversational goals from recognized key words and phrases within user input. Additionally, Seneff et al.’s <i>Galaxy-II: A Reference Architecture for Conversational System Development</i> (1998) describes conversational systems that parse recognized speech input to identify an appropriate domain and dialogue context based on lexical and syntactic cues.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>The intrinsic record is in accord. For example, the intrinsic record describes how:</p> <p>“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” ('699 patent at 16:37-42).</p> <p>Freeman and Baldwin (2006) disclose how cooperative conversation concepts were first described by Paul Grice in 1975 and how human-to-computer speech should ideally “follow the same model of human-to-human conversation” described in <i>Logic and Conversation</i>. Freeman and Baldwin explain:</p> <p>"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been."</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that the concept of using short-term and long-term shared knowledge to determine a context for an utterance was well understood, routine, and conventional:</p> <p>Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right? A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir. Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea? A. Correct.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|---|
| | <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Further, there is no technical disclosure or explanation in the specification about how a system specifically determines a context beyond the broad assertion that it leverages shared knowledge to model human conversation. The patent does not provide a specific algorithm, process, or computational method for identifying, distinguishing, or structuring context in a way that was not already known in the field. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of identifying a context for a natural language utterance based on recognized words or phrases was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[1D] determining, by the computer system, an interpretation of the natural language utterance based on the identified context;</p> | <p>The claimed method of determining, by the computer system, an interpretation of the natural language utterance based on the identified context was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to determine an interpretation of a natural language utterance from an identified context beyond broad conceptual statements. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Interpreting user utterances based on previously identified conversational or domain context was a well-established practice in dialogue system design. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes dialogue systems that interpret natural language input by incorporating the recognized</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>conversational context, such as domain or dialogue state. Additionally, Seneff et al.’s <i>Galaxy-II: A Reference Architecture for Conversational System Development</i> (1998) discusses conversational systems that determine the meaning of user input by interpreting recognized words and phrases in the context of the currently active domain model or dialogue frame.</p> <p>For example, the intrinsic record similarly describes</p> <p>“The utterance component of input 105 may be processed by a speech recognition engine 110 (which may alternatively be referred to herein as Automatic Speech Recognizer 110, or as shown in FIG. 1, ASR 110) to generate one or more preliminary interpretations of the utterance.” (’699 patent at 7:50-56).</p> <p>Additionally, the specification confirms that:</p> <p>“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>Freeman and Baldwin (2006) disclose how cooperative conversation concepts were first described by Paul Grice in 1975 and how human-to-computer speech should ideally “follow the same model of human-to-human conversation” described in <i>Logic and Conversation</i>. Freeman and Baldwin explain that intent determination, or meaning, is derived largely from context:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>"Intent determination, or meaning, is derived in a large part from the context behind the person's request."</p> <p>Tom Freeman and Larry Baldwin, Enhancing the VUE™ (Voice-User-Experience) Through Conversational Speech, SAE Technical Paper Series (2006), at 5.</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that the concept of using short-term and long-term shared knowledge to determine an interpretation for an utterance was well understood, routine, and conventional:</p> <p>Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right?</p> <p>A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir.</p> <p>Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea?</p> <p>A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Further, there is no technical disclosure or explanation in the specification about how a system specifically determines an interpretation beyond the broad assertion that it leverages shared knowledge to model human conversation. The patent does not provide a specific algorithm, process, or computational method for resolving ambiguities, prioritizing different contextual signals, or distinguishing between</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| | <p>multiple possible interpretations. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of determining an interpretation of a natural language utterance based on context was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[1E] accumulating, by the computer system, short-term knowledge based on one or more natural language utterances received during a predetermined time period, wherein the one or more natural language utterances received during the predetermined time period are related to a single conversation between a user and the computer system;</p> | <p>The claimed method of accumulating, by the computer system, short-term knowledge based on one or more natural language utterances received during a predetermined time period, wherein the utterances are related to a single conversation between a user and the computer system, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically accumulate short-term knowledge from temporally related utterances beyond general conceptual descriptions. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Maintaining and updating short-term conversational knowledge across multiple user utterances within a single dialogue session was a well-established practice. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes spoken dialogue systems that accumulate dialogue state information over sequences of related user inputs to maintain coherence and continuity during an interaction.</p> <p>Likewise, the intrinsic record describes how this reflected a well-understood and routine practice:</p> <p>“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>Additionally, the specification explains that:</p> <p>“Because cooperative conversations model human conversations, short-term session data may be expired after a psychologically appropriate amount of time, thereby humanizing system behavior, reducing a likelihood of contextual confusion based on stale data, while also adding relevant information from an expired session context to long-term knowledge models.” (’699 patent at 5:23-29).</p> <p>Freeman and Baldwin (2006) disclose how cooperative conversation concepts were first described by Paul Grice in 1975 and how human-to-computer speech should ideally “follow the same model of human-to-human conversation” described in Logic and Conversation. Freeman and Baldwin explain:</p> <p>"There is shared short-term (session) knowledge about what the user has said, what they have done, and where they are."</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that the concept of using short-term shared knowledge to track related utterances within a conversation was well understood, routine, and conventional:</p> <p>Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right? A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|---|
| | <p>Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea? A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Further, there is no technical disclosure or explanation in the specification about how a system specifically determines which utterances belong to a single conversation beyond the broad assertion that conversations should be handled similarly to human interactions. The patent does not provide a specific algorithm, process, or computational method for segmenting utterances into discrete conversational sessions or managing session data. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of accumulating short-term knowledge based on utterances received during a predetermined time period and associating them with a single conversation was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[1F] accumulating, by the computer system, long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period;</p> | <p>The claimed method of accumulating, by the computer system, long-term knowledge based on one or more natural language utterances received prior to the predetermined time period was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically accumulate long-term knowledge from prior utterances beyond broadly describing the concept. This lack of technical detail indicates that these</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Building and maintaining long-term conversational knowledge based on previously received utterances was a well-established practice in dialogue system and user modeling design. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes dialogue systems that maintain persistent user models and historical conversational data across sessions to enhance personalized interactions. Additionally, Kobsa’s <i>User Modeling and User-Adapted Interaction</i> (1993) discusses techniques for deploying user profiles in such systems.</p> <p>Likewise, the intrinsic record describes how:</p> <p style="padding-left: 40px;">“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>Additionally, the specification explains that:</p> <p style="padding-left: 40px;">“Because cooperative conversations model human conversations, short-term session data may be expired after a psychologically appropriate amount of time, thereby humanizing system behavior, reducing a likelihood of contextual confusion based on stale data, while also adding relevant information from an expired session context to long-term knowledge models.” (’699 patent at 5:23-29).</p> <p>Freeman and Baldwin (2006) disclose how cooperative conversation concepts were first described by Paul Grice in 1975 and how human-to-computer speech</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>should ideally “follow the same model of human-to-human conversation” described in Logic and Conversation. Freeman and Baldwin explain:</p> <p style="padding-left: 40px;">"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been."</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that the concept of long-term shared knowledge accumulation was well understood, routine, and conventional:</p> <p style="padding-left: 40px;">Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right? A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir. Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea? A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Further, there is no technical disclosure or explanation in the specification about how long-term knowledge is accumulated beyond the general notion that expired session data may be incorporated into long-term storage. The patent does not provide a specific algorithm, threshold, or computational method for determining which short-term knowledge should transition into long-term knowledge or how this differs from conventional knowledge management techniques. This</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| | <p>demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of accumulating long-term knowledge based on prior natural language utterances was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[2] The method of claim 1, wherein the manner in which the natural language utterance was spoken includes an indication of at least one of tone, pace, timing, inflection, word use, and/or jargon.</p> | <p>The claimed method of determining that the manner in which the natural language utterance was spoken includes an indication of at least one of tone, pace, timing, inflection, word use, and/or jargon was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically identify or interpret tone, pace, timing, inflection, word use, or jargon from a natural language utterance beyond broad conceptual statements. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Analyzing tone, speaking pace, timing, inflection patterns, and language use to infer characteristics about the user or conversation was a well-established practice in speech processing and dialogue systems. For example, Hirschberg’s <i>Communication and Prosody: Functional Aspects of Prosody</i> (1999) discusses how prosodic features such as pitch range, speech rate, and intonation contour contribute to conveying speaker attitudes, emotional state, and discourse structure. Additionally, Wahlster’s <i>SmartKom: Foundations of Multimodal Dialogue Systems</i> (2006) describes systems that detect and interpret prosodic and linguistic markers to adapt system responses dynamically.</p> <p>Likewise, the intrinsic record describes how:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>The ’699 patent explicitly discusses adapting responses based on these very factors but fails to disclose how such adaptation is technically achieved:</p> <p>“According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user’s manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response). Adapting to a user’s manner of speaking may include using contextual signifiers and grammatical rules to generate one or more sentences for use as response sets that may cooperate with the user.” (’699 patent at 18:45-54).</p> <p>While this passage describes desired outcomes, it does not explain how the system detects or processes tone, pace, timing, inflection, or word use, nor does it describe any specific algorithm or method for incorporating those elements into a response. Instead, it generically states that responses “may” be adapted, without disclosing any novel technological advancement that would distinguish the claimed method from conventional approaches.</p> <p>Freeman and Baldwin (2006) describe how conversational speech systems rely on shared knowledge, which includes various elements beyond just tone or manner of speaking:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been."</p> <p>In a conversational speech system both the system and the user share assumptions and expectations such as:</p> <p>Topic knowledge</p> <p>Conversation history</p> <p>Word usage and jargon</p> <p>Tone (e.g., formal, humorous, terse)</p> <p>The ability to make corrections without interrupting flow."</p> <p>This passage makes clear that tone and manner of speaking are just one of many aspects of shared knowledge used in conversation processing. The claimed limitation does not provide any disclosure that distinguishes tone, pace, or jargon from these other known aspects of shared conversational knowledge, nor does it describe any technical improvement in how tone or manner is detected and used. Instead, it broadly asserts that these characteristics may be taken into account without explaining how.</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that the concept of using shared knowledge—including tone, jargon, and conversation history—to interpret user input and generate responses was well understood, routine, and conventional:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| | <p>Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right?</p> <p>A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir.</p> <p>Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea?</p> <p>A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Further, there is no technical disclosure or explanation in the specification about how the system detects and processes the specified speech characteristics beyond the broad assertion that responses should reflect a conversational feel. The patent does not provide a specific algorithm, process, or computational method for analyzing tone, pace, inflection, or other vocal attributes in a way that was not already known. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of analyzing and incorporating tone, pace, timing, inflection, word use, or jargon into a natural language processing system was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[3] The method of claim 1, wherein the response comprises a voice response, and wherein generating the voice response based on the</p> | <p>The claimed method of generating a voice response based on the identified manner of speaking, including varying one or more of tone, pace, timing, inflection, word</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| <p>identified manner comprises varying one or more of tone, pace, timing, inflection, word use, and/or jargon of the voice response.</p> | <p>use, and/or jargon of the voice response, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to vary tone, pace, timing, inflection, word use, or jargon in generated voice responses beyond high-level conceptual statements. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Varying prosodic and linguistic features in system-generated speech to reflect conversational style or user-specific characteristics was a well-established technique. For example, Hirschberg’s <i>Communication and Prosody: Functional Aspects of Prosody</i> (1999) discusses how variations in pitch range, speech rate, and intonation can convey different speaker attitudes, emotional tones, or conversational nuances, suggesting techniques applicable to synthesized responses. Additionally, Wahlster’s <i>SmartKom: Foundations of Multimodal Dialogue Systems</i> (2006) describes systems that adjust the prosody and linguistic style of generated responses dynamically to enhance naturalness, engagement, and context-appropriate interaction.</p> <p>Likewise, the intrinsic record describes how:</p> <p>“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>The '699 patent explicitly discusses adapting responses based on these characteristics but does not disclose how this adaptation is technically accomplished:</p> <p>“According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user's manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response). Adapting to a user's manner of speaking may include using contextual signifiers and grammatical rules to generate one or more sentences for use as response sets that may cooperate with the user.” ('699 patent at 18:45-54).</p> <p>While this passage describes desired functionality, it does not explain how the system detects or processes tone, pace, timing, inflection, or word use, nor does it describe any specific algorithm or method for incorporating those elements into a response. Instead, it generically states that responses “may” be adapted, without disclosing any novel technological advancement that would distinguish the claimed method from conventional approaches.</p> <p>Freeman and Baldwin (2006) describe how conversational speech systems rely on shared knowledge, which includes elements beyond just tone or manner of speaking:</p> <p>"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been.”</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>In a conversational speech system both the system and the user share assumptions and expectations such as:</p> <p>Topic knowledge</p> <p>Conversation history</p> <p>Word usage and jargon</p> <p>Tone (e.g., formal, humorous, terse)</p> <p>The ability to make corrections without interrupting flow."</p> <p>This passage demonstrates that tone, pace, and manner of speech are just one of many aspects of shared knowledge that systems have long considered in generating responses. The claimed limitation does not distinguish tone, pace, or jargon from these other well-known aspects of shared conversational knowledge, nor does it describe any technical improvement in how those elements are identified and used in generating speech output.</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that the concept of adapting responses based on user speech characteristics was well understood, routine, and conventional:</p> <p>Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right?</p> <p>A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|---|
| | <p>Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea? A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Further, there is no technical disclosure or explanation in the specification about how the system generates a voice response that varies tone, pace, inflection, or word use beyond the broad assertion that responses should reflect a conversational feel. The patent does not provide a specific algorithm, process, or computational method for analyzing, modifying, or structuring voice responses in a way that was not already known. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of generating a voice response by varying tone, pace, timing, inflection, word use, or jargon was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[4pre] The method of claim 1, the method further comprising:</p> | |
| <p>[4A] obtaining, by the computer system, contextual signifiers and/or grammatical rules,</p> | <p>The claimed method of obtaining, by the computer system, contextual signifiers and/or grammatical rules was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically obtain contextual signifiers or grammatical rules beyond general statements. This lack of technical detail indicates that these processes were well</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Obtaining contextual signifiers and applying grammatical rules to assist in language understanding or generation was a well-established practice. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes dialogue systems that extract linguistic and contextual cues from user utterances to support natural language understanding and dialogue management. Additionally, Jurafsky’s <i>Lexical, Prosodic, and Syntactic Cues for Dialog Acts</i> (1998) discusses how recognized lexical items, prosodic features, and syntactic patterns serve as signifiers for identifying dialogue context and structure in conversational systems.</p> <p>Likewise, the intrinsic record describes how:</p> <p style="padding-left: 40px;">“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>The ’699 patent references contextual signifiers and grammatical rules but does not disclose how they are obtained or implemented in a meaningful way:</p> <p style="padding-left: 40px;">“By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for missing contextual signifiers, qualifiers, or other information. In one implementation, the response may ask for missing information in a way that most restricts possible interpretations, and the response may be</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>framed to establish a domain for a subsequent user utterance. In another implementation, common alternatives for nouns and verbs may be recognized to reflect variations in usage patterns according to various criteria.” ('699 patent at 3:11-22).</p> <p>This passage describes a desired outcome—that the system may model contextual signifiers or recognize grammatical variations—but does not disclose any method for actually obtaining or implementing contextual signifiers or grammatical rules. There is no algorithm, no defined structure, and no clear explanation of how the system would determine which grammatical rules to apply in a given context.</p> <p>Freeman and Baldwin (2006) discuss how shared conversational knowledge includes grammatical conventions and user-specific language patterns, but they place grammatical rules alongside a broad set of other known contextual factors:</p> <p style="padding-left: 40px;">"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been."</p> <p>In a conversational speech system both the system and the user share assumptions and expectations such as:</p> <ul style="list-style-type: none"> Topic knowledge Conversation history Word usage and jargon Tone (e.g., formal, humorous, terse) The ability to make corrections without interrupting flow." <p>This passage further demonstrates that contextual signifiers and grammatical rules are just one part of a broader framework of shared conversational knowledge and that they have long been used to process natural language in speech systems.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>Moreover, under oath, the inventors of the '699 patent admitted that using shared knowledge—including grammatical structures and contextual signifiers—to interpret user input was well understood, routine, and conventional:</p> <p>Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right?</p> <p>A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir.</p> <p>Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea?</p> <p>A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Furthermore, the '699 patent is indifferent to the type of computer being used, confirming that it relies on generic computing principles rather than any specialized implementation:</p> <p>“As used herein, a ‘request’ may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action.” ('699 patent at 2:32-35).</p> <p>This confirms that the patent does not claim any specific computing architecture or novel processing technique for obtaining contextual signifiers or grammatical rules.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| | <p>Further, there is no technical disclosure or explanation in the specification about how the system actually obtains contextual signifiers or grammatical rules beyond the broad assertion that they are modeled. The patent does not provide a specific algorithm, process, or computational method for retrieving, structuring, or applying grammatical rules in a way that was not already well known. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of obtaining contextual signifiers and grammatical rules was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[4B] wherein generating the response based on the identified manner in which the natural language utterance was spoken comprises using the obtained contextual signifiers and/or grammatical rules to generate sentences for use as response sets to cooperate with the user.</p> | <p>The claimed method of generating the response based on the identified manner in which the natural language utterance was spoken, including using the obtained contextual signifiers and/or grammatical rules to generate sentences for use as response sets to cooperate with the user, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically use contextual signifiers and grammatical rules to generate cooperative response sets beyond high-level descriptions. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Generating system responses by leveraging recognized contextual cues and grammatical structures to align with user input and maintain conversational cooperation was a well-established technique. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes dialogue systems that generate adaptive responses by applying grammatical rules and interpreting contextual signals from the user’s speech. Additionally, Jurafsky’s <i>Lexical, Prosodic, and Syntactic Cues</i></p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p><i>for Dialog Acts</i> (1998) discusses how systems can use lexical and syntactic patterns to shape the structure of system responses, supporting coherent, cooperative dialogue interactions.</p> <p>Likewise, the intrinsic record describes how:</p> <p>“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>The ’699 patent broadly states that responses may be generated in a way that cooperates with the user, but it does not disclose how this is actually accomplished:</p> <p>“According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user’s manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response). Adapting to a user’s manner of speaking may include using contextual signifiers and grammatical rules to generate one or more sentences for use as response sets that may cooperate with the user.” (’699 patent at 18:45-54).</p> <p>This passage describes a desired outcome—that responses “may” be structured using contextual signifiers and grammatical rules—but it fails to disclose any specific method, algorithm, or implementation details for how the system actually</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>generates responses in a cooperative manner. There is no explanation of how context is analyzed, how grammatical rules are applied dynamically, or how response sets are selected and structured.</p> <p>Additionally, the '699 patent confirms that it does not explain how contextual signifiers or grammatical rules are used in a technical sense:</p> <p style="padding-left: 40px;">“By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for missing contextual signifiers, qualifiers, or other information. In one implementation, the response may ask for missing information in a way that most restricts possible interpretations, and the response may be framed to establish a domain for a subsequent user utterance. In another implementation, common alternatives for nouns and verbs may be recognized to reflect variations in usage patterns according to various criteria.” ('699 patent at 3:11-22).</p> <p>This passage again describes functional goals rather than a specific technical implementation, confirming that the patent does not disclose a novel method for structuring responses using contextual signifiers or grammatical rules.</p> <p>Freeman and Baldwin (2006) discuss how conversational systems rely on shared knowledge, which includes both contextual signifiers and grammatical conventions, but they place these alongside many other well-known elements of conversation processing:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been. In a conversational speech system both the system and the user share assumptions and expectations such as:</p> <p>Topic knowledge Conversation history Word usage and jargon Tone (e.g., formal, humorous, terse) The ability to make corrections without interrupting flow." This passage makes clear that the use of contextual signifiers and grammatical structures to guide responses is just one part of a broader framework of shared conversational knowledge that has long been used in speech processing. The claimed limitation does not describe any novel improvement in how those elements are applied to generate response sets.</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that using shared knowledge—including grammatical structures and contextual signifiers—to interpret user input and generate responses was well understood, routine, and conventional:</p> <p>Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right? A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir. Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea?</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Furthermore, the '699 patent is indifferent to the type of computer being used, confirming that it relies on generic computing principles rather than any specialized implementation:</p> <p style="padding-left: 40px;">“As used herein, a ‘request’ may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action.” ('699 patent at 2:32-35).</p> <p>This confirms that the patent does not claim any specific computing architecture or novel processing technique for structuring responses.</p> <p>Further, there is no technical disclosure or explanation in the specification about how the system actually structures responses using contextual signifiers or grammatical rules beyond the broad assertion that response sets “may” be generated. The patent does not provide a specific algorithm, process, or computational method for analyzing, selecting, or applying grammatical rules in a way that was not already well known. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of generating response sets using contextual signifiers and grammatical rules was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| <p>[5pre] The method of claim 1, wherein the long-term knowledge is associated with a first user, the method further comprising:</p> | <p>The claimed method of associating the long-term knowledge with a first user, and further comprising additional steps based on that association, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically associate long-term knowledge with a particular user beyond broadly describing the concept. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Associating accumulated long-term knowledge with individual users to support personalized interactions and adaptive system behavior was a well-established practice in dialogue systems and user modeling. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes maintaining user-specific profiles that store user-specific data. Additionally, Kobsa’s <i>User Modeling and User-Adapted Interaction</i> (1993) discusses linking long-term knowledge repositories to individual users to enable adaptive spoken dialog systems.</p> |
| <p>[5A] generating, by the computer system, a profile associated with the first user based on the long-term knowledge, wherein the context for the natural language utterance is determined based further on the profile associated with the first user.</p> | <p>The claimed method of generating, by the computer system, a profile associated with the first user based on the long-term knowledge, wherein the context for the natural language utterance is determined further based on the user profile, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically generate a user profile from long-term knowledge or how to use that profile to further determine conversational context beyond broad conceptual descriptions. This lack of technical detail indicates that these processes were well</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Building user profiles from accumulated long-term knowledge and utilizing those profiles to inform conversational context determination was a well-established practice. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes dialogue systems that create persistent user profiles that aid in interpreting new utterances. Additionally, Kobsa’s <i>User Modeling and User-Adapted Interaction</i> (1993) discusses generating and maintaining structured user profiles that facilitate spoken dialog systems.</p> <p>.</p> <p>Likewise, the intrinsic record describes how:</p> <p style="padding-left: 40px;">“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>Additionally, the specification confirms that:</p> <p style="padding-left: 40px;">“Long-term shared knowledge may come from several sources, including user preferences or a plug-in data source (e.g., a subscription interface to a remote database), expertise of a user (e.g., based on a frequency of errors, types of tasks requested, etc., the user may be identified as a novice, intermediate, experienced, or other type of user), agent-specific information</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>and/or language that may also apply to other agents (e.g., by decoupling information from an agent to incorporate the information into other agents), frequently used topics passed in from the Session Input Accumulator, frequently used verbs, nouns, or other parts of speech, and/or other syntax information passed in from the Session Input Accumulator, or other sources of long-term shared knowledge may be used.” ('699 patent at 17:29-43).</p> <p>Freeman and Baldwin (2006) disclose how shared knowledge, including user history, is fundamental to conversation processing and adaptation:</p> <p style="padding-left: 40px;">"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been."</p> <p>This passage makes clear that user profiles, built from historical knowledge, were well known and commonly used to inform conversational context.</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that the concept of structuring user interactions based on short-term and long-term shared knowledge was well understood, routine, and conventional:</p> <p style="padding-left: 40px;">Q. And so the notion of -- the importance of using long-term and short-term shared knowledge to understand what is being said in a conversation is something that you got from Mr. Grice; right? A. This was part of the intellectual averment, and this was a very useful way of sharpening our vocabulary around this, the answer is yes, sir.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>Q. I just want to be clear, this was Mr. Grice’s idea, that wasn’t your idea? A. Correct.</p> <p>(Trial Tr. 289:4-16, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>The '699 patent describes the use of adaptive responses based on user-specific factors, but it fails to explain how user profiles are actually generated or structured:</p> <p>“According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user's manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response).” ('699 patent at 18:45-54).</p> <p>This passage merely describes desired outcomes—that responses "may adapt" based on user factors—but does not explain how user profiles are created, updated, or used to determine context.</p> <p>Additionally, the '699 patent fails to provide any meaningful technical disclosure regarding how user profiles are integrated into conversational context determination:</p> <p>“By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>missing contextual signifiers, qualifiers, or other information.” (’699 patent at 3:11-22).</p> <p>This passage confirms that the patent does not disclose a method for structuring or applying user profiles to conversational context processing. Instead, it generically states that context may be adjusted without explaining how user-specific factors are incorporated into the process.</p> <p>Furthermore, the ’699 patent is indifferent to the type of computer being used, confirming that it relies on a generic computing architecture rather than any specialized processing technique:</p> <p>“As used herein, a ‘request’ may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action.” (’699 patent at 2:32-35).</p> <p>This confirms that the patent does not claim any specific computing architecture or novel processing technique for generating or applying user profiles.</p> <p>Further, there is no technical disclosure or explanation in the specification about how the system actually generates and applies a user profile to determine conversational context beyond the broad assertion that user-specific factors “may” influence responses. The patent does not provide a specific algorithm, process, or computational method for defining, ranking, or integrating user profiles in a way that was not already well known. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| | Thus, the claimed method of generating a user profile based on long-term knowledge and using that profile to determine conversational context was well understood, routine, and conventional in the relevant industry at the time of filing. |
| [6pre] The method of claim 1, wherein generating the response based on the identified manner comprises: | |
| [6A] adapting, by the computer system, the response based on a response format associated with the identified manner. | <p>The claimed method of adapting, by the computer system, the response based on a response format associated with the identified manner of speaking was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically adapt a system response based on an identified manner of speaking and a corresponding response format beyond general conceptual statements. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Adapting system responses to align with an identified speaking style, prosodic pattern, or conversational manner using pre-associated response formats was a well-established technique. For example, Hirschberg’s <i>Communication and Prosody: Functional Aspects of Prosody</i> (1999) discusses how variations in user prosody—such as tone, pace, and inflection—can facilitate spoken dialog systems. Additionally, Wahlster’s <i>SmartKom: Foundations of Multimodal Dialogue Systems</i> (2006) describes dialogue systems that adjust the structure, style, and prosodic features of system responses based on the detected manner of user interaction.</p> <p>Likewise, the intrinsic record describes how:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>“Just as humans rely on knowledge shared by participants to examine how much and what kind of information was available, Intelligent Hypothesis Builder 310 may leverage the identified conversation type and short-term and long-term shared knowledge to generate a degree of certainty for each hypothesis.” (’699 patent at 16:37-42).</p> <p>The ’699 patent discusses adaptive responses and structuring responses based on user characteristics but fails to disclose how this adaptation is achieved:</p> <p>“According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user’s manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response).” (’699 patent at 18:45-54).</p> <p>While this passage describes an adaptive response, it does not provide any technical detail on how a response format is determined, structured, or varied based on user characteristics. Instead, it broadly asserts that responses may be adapted without explaining any method for accomplishing this.</p> <p>Freeman and Baldwin (2006) disclose the adaptive response framework, which describes how conversational responses should be determined:</p> <p>Adaptive Responses</p> <p>"As in a human conversation, the response of the system should be based on how confident the system is that its conclusion is correct. If it is</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>certain, it should just execute the request. If it is less than certain but the range of choice is few or none, it should execute. If it is less than certain and the choices are many, it should ask."</p> <p>This framework demonstrates that the concept of adapting responses based on confidence and conversational context was well known at the relevant time.</p> <p>Moreover, under oath, Mr. Freeman admitted that the adaptive response concept described in Freeman and Baldwin (2006) corresponds to the limitations in the claims, where the '681 patent is the parent of the '699 patent:</p> <p>Q. And this marketing paper where you describe these principles that Mr. Grice had laid out, this is essentially the '681 patent in ordinary language, right? A. So this is '681 but '681 takes this to a machine, as opposed to human to human. (Trial Tr. 289:20-24, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Furthermore, Mr. Freeman confirmed that adaptive response was derived from Mr. Grice’s work:</p> <p>Q. And adaptive response, that also came from Mr. Grice's paper? A. Correct.</p> <p>(Trial Tr. 289:17-19, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>Additionally, the '699 patent provides no meaningful technical disclosure regarding how response formats are structured, determined, or applied:</p> <p>“By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for missing contextual signifiers, qualifiers, or other information.” ('699 patent at 3:11-22).</p> <p>This passage confirms that the patent does not disclose a method for structuring a response format beyond the broad assertion that responses may be adapted based on user characteristics. There is no technical detail explaining how responses are framed, what data structures are used, or how responses are evaluated for appropriateness.</p> <p>Freeman and Baldwin (2006) explain that adaptation in speech systems is just one part of a broader set of shared knowledge principles:e</p> <p>"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been.</p> <p>In a conversational speech system both the system and the user share assumptions and expectations such as:</p> <p>Topic knowledge</p> <p>Conversation history</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| | <p>Word usage and jargon</p> <p>Tone (e.g., formal, humorous, terse)</p> <p>The ability to make corrections without interrupting flow."</p> <p>This passage confirms that response adaptation based on format is just one of many aspects of shared conversational knowledge, and there is nothing novel about using shared knowledge—including tone, jargon, and speech style—to influence response structures.</p> <p>Further, there is no technical disclosure or explanation in the specification about how the system determines, selects, or applies a response format beyond the broad assertion that responses may be structured according to user mannerisms. The patent does not provide a specific algorithm, process, or computational method for defining, ranking, or integrating response formatting rules in a way that was not already well known. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of adapting a response based on a response format associated with a user’s identified manner was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[7pre] The method of claim 1, wherein recognizing the one or more words or phrases from the natural language utterance comprises:</p> | |
| <p>[7A] providing, by the computer system, the natural language utterance as an input to a speech recognition engine; and</p> | <p>The claimed method of providing, by the computer system, the natural language utterance as an input to a speech recognition engine was well understood, routine, and conventional in the relevant industry at the relevant time.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>There is no technical disclosure or explanation in the specification about how to specifically provide a natural language utterance to a speech recognition engine beyond a general conceptual description. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Providing user speech input to a speech recognition engine for processing was a fundamental and well-established practice in natural language and dialogue systems. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes numerous dialogue systems that route user utterances directly into speech recognition components to generate recognized word sequences for further processing.</p> <p>Moreover, the '699 patent is indifferent to the type of computer performing this function, confirming that it relies on a general-purpose computing system rather than any specialized or improved computing architecture:</p> <p style="padding-left: 40px;">“As used herein, a ‘request’ may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action.” ('699 patent at 2:32-35).</p> <p>The intrinsic record further describes how utterances are processed by generic speech recognition technology:</p> <p style="padding-left: 40px;">“The utterance component of input 105 may be processed by a speech recognition engine 110 (which may alternatively be referred to herein as Automatic Speech Recognizer 110, or as shown in FIG. 1, ASR 110)</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>to generate one or more preliminary interpretations of the utterance.” ('699 patent at 7:50-56).</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that automatic speech recognition (ASR) engines were commodity products at the time of filing:</p> <p>Q. And that software section, let’s go to the next page, 0194. And if we could on that page, if you could put just the top paragraph on the screen. And kind of put it to the side so you can see both at the same time. During the opening by Amazon’s counsel, you indicated that the original Cybermind used some software from IBM, did you hear that?</p> <p>A. Yes.</p> <p>Q. What was that, sir?</p> <p>A. That was a component called the ASR, Automated Speech Recognition engine, and that was pretty much a commodity product at the same time.</p> <p>(Trial Tr. 147:18-148:3, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>This confirms that ASR systems were well known and widely available at the relevant time, and that the claimed computer system does nothing beyond what was already standard practice.</p> <p>Furthermore, the '699 patent does not disclose any improvements or novel techniques related to speech recognition but instead assumes a pre-existing, conventional speech recognition system:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| | <p>“By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for missing contextual signifiers, qualifiers, or other information.” (’699 patent at 3:11-22).</p> <p>This passage confirms that the patent does not claim any innovation in how utterances are processed by an ASR engine.</p> <p>Thus, the claimed method of providing a natural language utterance as an input to a speech recognition engine was well understood, routine, and conventional in the relevant industry, and the claimed system relies entirely on generic computing components performing known functions.</p> |
| <p>[7B] obtaining, by the computer system, the one or more words or phrases recognized from the natural language utterance as an output of the speech recognition engine.</p> | <p>The claimed method of obtaining, by the computer system, one or more words or phrases recognized from the natural language utterance as an output of the speech recognition engine was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically obtain recognized words or phrases from a speech recognition engine beyond a general conceptual description. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Obtaining recognized word sequences or phrases from a speech recognition engine after processing a natural language utterance was a well-established step in speech-enabled dialogue systems. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes systems where the output of the speech recognizer—</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>recognized words and phrases—is passed forward for interpretation and dialogue management.</p> <p>The concept of obtaining recognized words or phrases as an output from a speech recognition engine was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>The '699 patent is indifferent to the type of computer performing this function, confirming that it relies on a generic computing system rather than any specialized or improved computing architecture:</p> <p style="padding-left: 40px;">“As used herein, a ‘request’ may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action.” ('699 patent at 2:32-35).</p> <p>The intrinsic record further describes how generic speech recognition technology generates recognized words or phrases as an output:</p> <p style="padding-left: 40px;">“The utterance component of input 105 may be processed by a speech recognition engine 110 (which may alternatively be referred to herein as Automatic Speech Recognizer 110, or as shown in FIG. 1, ASR 110) to generate one or more preliminary interpretations of the utterance.” ('699 patent at 7:50-56).</p> <p>Moreover, under oath, the inventors of the '699 patent admitted that automatic speech recognition (ASR) engines were well known and widely available at the time of filing:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>Q. And that software section, let’s go to the next page, 0194. And if we could on that page, if you could put just the top paragraph on the screen. And kind of put it to the side so you can see both at the same time. During the opening by Amazon’s counsel, you indicated that the original Cybermind used some software from IBM, did you hear that?</p> <p>A. Yes.</p> <p>Q. What was that, sir?</p> <p>A. That was a component called the ASR, Automated Speech Recognition engine, and that was pretty much a commodity product at the same time.</p> <p>(Trial Tr. 147:18-148:3, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>This confirms that ASR systems were well known, widely used, and performed the exact function described in the claim—receiving a spoken utterance and outputting recognized words or phrases. The claim does not introduce any improvement to ASR processing, nor does it specify any novel method for extracting recognized words.</p> <p>Furthermore, the '699 patent does not disclose any innovations related to speech recognition output but assumes pre-existing ASR systems performing standard recognition tasks:</p> <p>“By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for missing contextual signifiers, qualifiers, or other information.” ('699 patent at 3:11-22).</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|---|
| | <p>This passage further confirms that the patent does not claim any improvement in how speech recognition engines generate recognized words or phrases as output.</p> <p>Thus, the claimed method of obtaining recognized words or phrases from a speech recognition engine was well understood, routine, and conventional in the relevant industry, and the claimed system relies entirely on generic computing components performing known functions.</p> |
| <p>[8pre] The method of claim 1, the method further comprising:</p> | |
| <p>[8A] causing, by the computer system, the response to the natural language utterance to be provided to the user.</p> | <p>The claimed method of causing, by the computer system, the response to the natural language utterance to be provided to the user was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically cause the response to be provided to the user beyond general conceptual descriptions. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Providing a generated system response to the user based on the interpretation of a natural language utterance was a fundamental and well-established practice. For example, McTear’s <i>Spoken Dialogue Technology</i> (2004) describes numerous dialogue systems that deliver system responses to users via output modalities such as speech synthesis or graphical displays following interpretation of user input.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>The '699 patent confirms that the claimed system is a generic computing system performing known functions rather than any specialized or improved computing architecture:</p> <p>“As used herein, a ‘request’ may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action.” ('699 patent at 2:32-35).</p> <p>The intrinsic record describes how generic systems generate responses to user utterances, e.g.:</p> <p>“System 100 may generate a response based on an interpretation of the user input and may present the response via an output device 115.</p> <p>Moreover, Freeman and Baldwin (2006) disclose the concept of an “adaptive response,” confirming that generating and providing a response based on user input was well known:</p> <p>Adaptive Responses</p> <p>"As in a human conversation, the response of the system should be based on how confident the system is that its conclusion is correct. If it is certain, it should just execute the request. If it is less than certain but the range of choice is few or none, it should execute. If it is less than certain and the choices are many, it should ask."</p> <p>Under oath, Mr. Freeman admitted that this “adaptive response” concept corresponds to the limitations in the claim:</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>Q. And this marketing paper where you describe these principles that Mr. Grice had laid out, this is essentially the '681 patent in ordinary language, right?</p> <p>A. So this is '681 but '681 takes this to a machine, as opposed to human to human. (Trial Tr. 289:20-24, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Furthermore, Mr. Freeman confirmed that adaptive response came directly from Grice’s work:</p> <p>Q. And adaptive response, that also came from Mr. Grice's paper?</p> <p>A. Correct.</p> <p>(Trial Tr. 289:17-19, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>The '699 patent describes the generation of responses but does not provide any meaningful technical disclosure on how this is done:</p> <p>“According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user's manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing, inflection, word use, jargon, and other variables in a verbal or audible response).” ('699 patent at 18:45-54).</p> <p>This passage does not disclose a technical method for generating responses or how responses are provided to a user but instead describes a goal—that responses “may” be generated in a certain way.</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|---|
| | <p>Furthermore, there is no technical disclosure or explanation in the specification about how responses are structured, processed, or delivered to the user beyond the broad assertion that responses are provided. The patent does not provide a specific algorithm, process, or computational method for improving response generation or output. This demonstrates that it was well understood, routine, and conventional in the relevant industry at the time.</p> <p>Thus, the claimed method of causing a response to be provided to a user was well understood, routine, and conventional in the relevant industry, and the claimed system relies entirely on generic computing components performing known functions.</p> |
| <p>9. The method of claim 1, wherein the response comprises a voice response, and wherein generating the voice response to the natural language utterance based on the identified manner comprises varying one or more of word use and/or jargon of the voice response.</p> | <p>The claimed method of generating a voice response to the natural language utterance based on the identified manner, including varying one or more of word use and/or jargon of the voice response, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically vary word use or jargon in the generated voice response based on the user's manner of speaking beyond high-level conceptual statements. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Adapting system-generated voice responses by adjusting word choice and jargon to align with user mannerisms or conversational style was a recognized and well-established practice in dialogue system design. For instance, Wahlster’s <i>SmartKom: Foundations of Multimodal Dialogue Systems</i> (2006) describes multimodal dialogue systems that adapt their language output based on user-</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>specific and contextual interaction factors to improve engagement and understanding.</p> <p>Likewise, the '699 patent confirms that the claimed system does not rely on any specialized or improved computing architecture but instead performs well-known functions on a generic computing system:</p> <p style="padding-left: 40px;">“As used herein, a ‘request’ may be a command, directive, or other instruction for a device, computer, or other machine to retrieve information, perform a task, or take some other action.” (‘699 patent at 2:32-35).</p> <p>The intrinsic record further describes how a system may generate responses:</p> <p style="padding-left: 40px;">“In one implementation, the intelligent responses may include a verbal or 20 audible reply played through an output device (e.g., a speaker), and/or an action performed by a device, computer, or machine (e.g., downloading a web page, showing a list, executing an application, etc.).” (‘765 patent at 18:18-23).</p> <p>The '699 patent acknowledges the concept of adapting responses based on user speech characteristics but fails to disclose how this adaptation is actually accomplished:</p> <p style="padding-left: 40px;">“According to another aspect of the invention, Adaptive Response Builder 315 may create a conversational feel by adapting to a user's manner of speaking, framing responses appropriately, and having natural variation and/or personality (e.g., by varying tone, pace, timing,</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|--|
| | <p>inflection, word use, jargon, and other variables in a verbal or audible response).” (’699 patent at 18:45-54).</p> <p>While this passage states that responses may be adapted based on user speech, it does not describe any specific method, algorithm, or technique for varying word use or jargon dynamically. Instead, it provides a broad, high-level assertion that responses “may” be adapted without explaining how.</p> <p>Freeman and Baldwin (2006) disclose the adaptive response framework, confirming that response adaptation was a well-known technique:</p> <p style="padding-left: 40px;">"As in a human conversation, the response of the system should be based on how confident the system is that its conclusion is correct. If it is certain, it should just execute the request. If it is less than certain but the range of choice is few or none, it should execute. If it is less than certain and the choices are many, it should ask."</p> <p>Under oath, Mr. Freeman admitted that this “adaptive response” concept corresponds to the limitations in the claim:</p> <p style="padding-left: 40px;">Q. And this marketing paper where you describe these principles that Mr. Grice had laid out, this is essentially the '681 patent in ordinary language, right?</p> <p style="padding-left: 40px;">A. So this is '681 but '681 takes this to a machine, as opposed to human to human.</p> <p style="padding-left: 40px;">(Trial Tr. 289:20-24, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p style="padding-left: 40px;">Furthermore, Mr. Freeman confirmed that adaptive response was derived from Mr. Grice’s work:</p> <p style="padding-left: 40px;">Q. And adaptive response, that also came from Mr. Grice's paper?</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>A. Correct.</p> <p>(Trial Tr. 289:17-19, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Freeman and Baldwin (2006) explain that adaptation in speech systems is just one part of a broader set of shared knowledge principles:</p> <p>"There is also long-term shared knowledge which is user-centric rather than session-based such as who they are, their preferences, history of what they've done and said, and where they have been.</p> <p>In a conversational speech system both the system and the user share assumptions and expectations such as:</p> <p>Topic knowledge</p> <p>Conversation history</p> <p>Word usage and jargon</p> <p>Tone (e.g., formal, humorous, terse)</p> <p>The ability to make corrections without interrupting flow."</p> <p>This passage confirms that varying word use and jargon in a response is just one of many aspects of shared conversational knowledge that has long been used in speech processing. The claimed limitation does not describe any novel</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| | <p>improvement in how those elements are identified and used in generating speech output.</p> <p>The '699 patent provides no meaningful technical disclosure regarding how response variation is determined or structured:</p> <p style="padding-left: 40px;">“By modeling what contextual signifiers, qualifiers, or other information may be required to perform a task in an identified context, an adaptive response may be generated, such as prompting a user for missing contextual signifiers, qualifiers, or other information.” ('699 patent at 3:11-22).</p> <p>This passage confirms that the patent does not disclose any technical details on how response variations—such as word choice or jargon—are implemented in practice.</p> <p>Further, there is no technical disclosure or explanation in the specification about how the system determines, selects, or applies variations in word use or jargon beyond the broad assertion that responses may be structured according to user mannerisms. The patent does not provide a specific algorithm, process, or computational method for defining, ranking, or integrating response formatting rules in a way that was not already well known.</p> <p>Thus, the claimed method of generating a voice response by varying word use and/or jargon was well understood, routine, and conventional in the relevant industry at the time of filing.</p> |
| <p>[10] The method of claim 1, the method further comprising:</p> | <p>The claimed method of adapting the response, by the computer system, to model a conversation, wherein adapting the response to model a conversation comprises</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| <p>adapting the response, by the computer system, to model a conversation, wherein adapting the response to model a conversation comprises adapting the response to have a personality by varying word use.</p> | <p>adapting the response to have a personality by varying word use, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically adapt a system response to model a conversation with a personality through varied word use beyond broad conceptual descriptions. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Modeling conversational personality by varying lexical choices to reflect different interaction styles was a well-established practice in dialogue system design. Additionally, Wahlster’s <i>SmartKom: Foundations of Multimodal Dialogue Systems</i> (2006) describes systems that adapt language use dynamically to create more natural, engaging, and personality-driven interactions with users.</p> <p>Freeman and Baldwin (2006) disclose the adaptive response framework, confirming that response adaptation was a well-known technique:</p> <p style="padding-left: 40px;">"As in a human conversation, the response of the system should be based on how confident the system is that its conclusion is correct. If it is certain, it should just execute the request. If it is less than certain but the range of choice is few or none, it should execute. If it is less than certain and the choices are many, it should ask."</p> <p>Under oath, Mr. Freeman admitted that this “adaptive response” concept corresponds to the limitations in the claim:</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|--|
| | <p>Q. And this marketing paper where you describe these principles that Mr. Grice had laid out, this is essentially the '681 patent in ordinary language, right?</p> <p>A. So this is '681 but '681 takes this to a machine, as opposed to human to human.</p> <p>(Trial Tr. 289:20-24, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> <p>Furthermore, Mr. Freeman confirmed that adaptive response was derived from Mr. Grice’s work:</p> <p>Q. And adaptive response, that also came from Mr. Grice's paper?</p> <p>A. Correct.</p> <p>(Trial Tr. 289:17-19, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)).</p> |
| <p>[11] The method of claim 1, the method further comprising generating a response that is sensitive to context, what the user already knows about a topic, short-term knowledge and long-term knowledge of user preferences, and words uttered by the user in one or more prior natural language utterances.</p> | <p>The claimed method of generating a response that is sensitive to context, what the user already knows about a topic, short-term knowledge and long-term knowledge of user preferences, and words uttered by the user in one or more prior natural language utterances, was well understood, routine, and conventional in the relevant industry at the relevant time.</p> <p>There is no technical disclosure or explanation in the specification about how to specifically generate such context-sensitive and user-adaptive responses beyond general conceptual statements. This lack of technical detail indicates that these processes were well understood, routine, and conventional in the relevant industry at the relevant time period.</p> <p>Generating responses sensitive to dialogue context, user background knowledge, conversational history, and user preferences was a well-established practice in</p> |

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|----------------|---|
| | <p>dialogue system and user modeling research. For example, McTear's <i>Spoken Dialogue Technology</i> (2004) describes dialogue systems that generate adaptive responses based on dialogue context, user interaction history, and inferred user goals. Additionally, Kobsa's <i>User Modeling and User-Adapted Interaction</i> (1993) discusses building user models that facilitate spoken dialog systems.</p> <p>Freeman and Baldwin (2006) disclose the adaptive response framework, confirming that response adaptation was a well-known technique:</p> <p style="padding-left: 40px;">"As in a human conversation, the response of the system should be based on how confident the system is that its conclusion is correct. If it is certain, it should just execute the request. If it is less than certain but the range of choice is few or none, it should execute. If it is less than certain and the choices are many, it should ask."</p> <p>Under oath, Mr. Freeman admitted that this "adaptive response" concept corresponds to the limitations in the claim:</p> <p style="padding-left: 40px;">Q. And this marketing paper where you describe these principles that Mr. Grice had laid out, this is essentially the '681 patent in ordinary language, right?</p> <p style="padding-left: 40px;">A. So this is '681 but '681 takes this to a machine, as opposed to human to human. (Trial Tr. 289:20-24, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)). Furthermore, Mr. Freeman confirmed that adaptive response was derived from Mr. Grice's work:</p> <p style="padding-left: 40px;">Q. And adaptive response, that also came from Mr. Grice's paper?</p> <p style="padding-left: 40px;">A. Correct.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|---|
| | (Trial Tr. 289:17-19, VB Assets v. Amazon.com Services LLC, No. 19-cv-01410, at ECF No. 304 (D. Del. Dec. 12, 2023)). |
| [12P] A system for generating natural language system responses adapted based on a user's manner of speaking, the system comprising: | Claim element 12P is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1P]. |
| [12A] one or more physical processors programmed with one or more computer program instructions which, when executed, configure the one or more physical processors to: | The Federal Circuit has long held that employing generic computer components does not save a claim from ineligibility. <i>See Voter Verified, Inc. v. Election Systems & Software, LLC</i> , 887 F.3d 1376, 1385, 126 USPQ2d 1498, 1504 (Fed. Cir. 2018). |
| [12B] receive a user input comprising a natural language utterance; | Claim element 12B is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1A]. |
| [12C] recognize one or more words or phrases from the natural language utterance; | Claim element 12C is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1B]. |
| [12D] identify a context for the natural language utterance based on the one or more words or phrases recognized from the natural language utterance; | Claim element 12D is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1D]. |
| [12E] determine an interpretation of the natural language utterance based on the identified context; | Claim element 12E is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1D]. |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| [12F] accumulate short-term knowledge based on one or more natural language utterances received during a predetermined time period, wherein the one or more natural language utterances received during the predetermined time period are related to a single conversation between a user and the computer system; | Claim element 12F is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1E]. |
| [12G] accumulate long-term knowledge, wherein the long-term knowledge is accumulated based on one or more natural language utterances received prior to the predetermined time period; | Claim element 12G is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1F]. |
| [12H] identify a manner in which the natural language utterance was spoken based on the short-term knowledge and the long-term knowledge; and | Claim element 12H is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1G]. |
| [12I] generate a response to the natural language utterance based on the interpretation and the identified manner in which the natural language utterance was spoken. | Claim element 12I is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [1H]. |
| [13] The system of claim 12, wherein the manner in which the natural language utterance was spoken includes an indication of at least one | Claim element 13 is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [2]. |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidity Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| of tone, pace, timing, inflection, word use, and/or jargon. | |
| [14] The system of claim 12, wherein the response comprises a voice response, and wherein to generate the voice response based on the identified manner, the one or more physical processors are further configured to: vary one or more of tone, pace, timing, inflection, word use, and/or jargon of the voice response. | Claim element 14 is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [3]. |
| [15A] The system of claim 12, wherein the one or more physical processors are further configured to: obtain contextual signifiers and/or grammatical rules, | Claim element 15A is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [4A]. |
| [15B] wherein to generate the response based on the identified manner in which the natural language utterance was spoken, the one or more physical processors are further configured to use the obtained contextual signifiers and/or grammatical rules to generate sentences for use as response sets to cooperate with the user. | Claim element 15B is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [4B]. |
| [16A] The system of claim 12, wherein the long-term knowledge is associated with a first user, and | Claim element 16A is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [5A]. |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|---|---|
| <p>[16B] wherein the one or more physical processors are further configured to:</p> <p>generate a profile associated with the first user based on the long-term knowledge, wherein the context for the natural language utterance is determined based further on the profile associate with the first user.</p> | <p>Claim element 16B is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [5B].</p> |
| <p>[17] The system of claim 12, wherein to generate the response based on the identified manner, the one or more physical processors are configured to:</p> <p>adapt the response based on a response format associated with the identified manner.</p> | <p>Claim element 17 is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [6].</p> |
| <p>[18A] The system of claim 12, wherein to recognize the one or more words or phrases from the natural language utterance, the one or more physical processors are configured to:</p> <p>provide the natural language utterance as an input to a speech recognition engine; and</p> | <p>Claim element 18A is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [7A].</p> |
| <p>[18B] obtain the one or more words or phrases recognized from the natural language utterance as an output of the speech recognition engine.</p> | <p>Claim element 18B is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [7B].</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidity Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|--|
| <p>[19] The system of claim 12, wherein the one or more physical processors are further configured to: cause the response to the natural language utterance to be provided to the user.</p> | <p>Claim element 19 is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [8].</p> |
| <p>[20] The system of claim 12, wherein the response comprises a voice response, and wherein to generate the voice response to the natural language utterance based on the identified manner, the one or more physical processors are further configured to: vary one or more of word use and/or jargon of the voice response.</p> | <p>Claim element 20 is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [9].</p> |
| <p>[21] The system of claim 12, wherein the one or more physical processors are further configured to: adapt the response to model a conversation, wherein adapting the response to model a conversation comprises adapting the response to have a personality by varying word use.</p> | <p>Claim element 21 is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [10].</p> |
| <p>[22] The system of claim 12, wherein the one or more physical processors are further configured to generate a response that is sensitive to context, what the user already knows about a topic, short-term knowledge and long-term</p> | <p>Claim element 22 is well-understood, routine, and conventional for at least the reasons set forth above. <i>See</i> claim element [11].</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants' Invalidation Contentions – Appendix G-3

| Claim Language | Description of the Industry and Description of How Each Element Was Well-Understood, Routine, and Conventional |
|--|---|
| knowledge of user preferences, and words uttered by the user in one or more prior natural language utterances. | |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

Pursuant to Judge Gilstrap’s Standing Order Regarding Subject Matter Eligibility Contentions, Samsung provides the following chart identifying any other factual or legal basis for why the Challenged Claims are otherwise ineligible for patent protection.

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|--|
| 1-22 | <p>The challenged claims reflect long-understood and well-known concepts in linguistics, computational linguistics, and natural language understanding, as reflected in the cited references, in addition to the following:</p> <p>Bub, T., Wahlster, W., & Waibel, A. <i>Verbmobil: The combination of deep and shallow processing for spontaneous speech translation</i>. German Research Center for Artificial Intelligence; Carnegie Mellon University.</p> <p>Oh, A. H., & Rudnicky, A. I.. <i>Stochastic language generation for spoken dialogue systems</i>. Carnegie Mellon University.</p> <p>Batliner, A., Zeißler, V., Frank, C., Adelhardt, J., Shi, R. P., & Nöth, E. (2003). <i>We are not amused – But how do you know? User states in a multi-modal dialogue system</i>. University of Erlangen-Nuremberg.</p> <p>Aust, H., Oerder, M., Seide, F., & Steinbiss, V. (1995). <i>A spoken language inquiry system for automatic train timetable information</i>. Philips Research Laboratories.</p> <p>Jönsson, A.. <i>A dialogue manager using initiative-response units and distributed control</i>. Linköping University.</p> <p>Aust, H., Oerder, M., Seide, F., & Steinbiss, V. (1994). <i>Experience with the Philips automatic train timetable information system</i>. Philips Research Laboratories.</p> <p>Juneja, A. (2004). <i>Approaches to dialogue system design</i>. Carnegie Mellon University.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|---|
| | <p>Allen, J. F., Byron, D. K., Dzikovska, M., Ferguson, G., Galescu, L., & Stent, A. (2001). Towards conversational human-computer interaction. <i>AI Magazine</i>, 22(4), 27–37.</p> <p>Gorin, A. L., Parker, B. A., Sachs, R. M., & Wilpon, J. G. (1996). How may I help you? AT&T Research.</p> <p>Allen, J., Byron, D., Dzikovska, M., Ferguson, G., Galescu, L., & Stent, A. (1998). An architecture for a generic dialogue shell. <i>Natural Language Engineering</i>, 1(1), 1–24.</p> <p>Veldhuijzen van Zanten, G. (1998). Adaptive mixed-initiative dialogue management. IPO Center for Research on User-System Interaction.</p> <p>Allen, J., Ferguson, G., & Stent, A. (2001). An architecture for more realistic conversational systems. University of Rochester.</p> <p>Banuazizi, A., & Creswell, C. (n.d.). Is that a real question?: Final rises, final falls and discourse function in yes-no question intonation. University of Pennsylvania.</p> <p>Bohus, D., Raux, A., Harris, T. K., Eskenazi, M., & Rudnicky, A. I. (2007). Olympus: An open-source framework for conversational spoken language interface research. Carnegie Mellon University.</p> <p>Blaylock, N. (2002). Managing communicative intentions in dialogue using a collaborative problem-solving model (Technical Report 774). University of Rochester.</p> <p>Microsoft. (2004). Microsoft Speech Server 2004: Product datasheet. Microsoft Corporation.</p> <p>Graesser, A. C., VanLehn, K., Rosé, C. P., Jordan, P. W., & Harter, D. (2001). Intelligent tutoring systems with conversational dialogue. <i>AI Magazine</i>, 22(4), 39–51.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|---|
| | <p>Hieronymus, J. L., & Dowding, J. (2004). Clarissa spoken dialogue system for procedure reading and navigation. NASA Ames Research Center.</p> <p>LumenVox.. LumenVox Speech Driven Assistant: Printed documentation. LumenVox.</p> <p>Walton, D. (2007). Dialogical models of explanation. Association for the Advancement of Artificial Intelligence.</p> <p>Cohen, P. R., Johnston, M., McGee, D., Oviatt, S., Pittman, J., Smith, I., Chen, L., & Glow, J. (1997). QuickSet: Multimodal interaction for distributed applications. Oregon Graduate Institute of Science and Technology.</p> <p>Cohen, P. R. (1997). Modeling speech acts for natural language interfaces. Oregon Graduate Institute of Science and Technology.</p> <p>Johnston, M. (2000, November 20). Comdex: One Voice shows off talking Web navigator 'TVAN'. Network World, IDG Communications, Inc.</p> <p>Ferguson, G., & Allen, J. F. (1998). TRIPS: An integrated intelligent problem-solving assistant. University of Rochester.</p> <p>Allen, J. F., Ferguson, G., & Stent, A. (2001). An architecture for more realistic conversational systems. University of Rochester.</p> <p>Nguyen, P. (2005, March). The first step: Open speech standards. Customer Interaction Solutions, 23(9), 1–2.</p> <p>Chai, J., Pan, S., Zhou, M. X., & Houck, K. (2005). Context-based multimodal input understanding in conversational systems. IBM T. J. Watson Research Center.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|---|
| | <p>Doherty, P., Haslum, P., Merz, T., Skarman, E., Conte, G., Duranti, S., Heintz, F., Pettersson, P.-O., Persson, T., & Wingman, B. (2004). A distributed architecture for intelligent unmanned aerial vehicle experimentation. Linköping University.</p> <p>Chu-Carroll, J. (2000). MIMIC: An adaptive mixed initiative spoken dialogue system for information queries. Lucent Technologies Bell Laboratories.</p> <p>Alexandersson, J., Maier, E., & Reithinger, N. (1995). A robust and efficient three-layered dialogue component for a speech-to-speech translation system. DFKI GmbH.</p> <p>Sitter, S., & Stein, A. (1992). Modeling information-seeking dialogues: The conversational roles (COR) model. GMD-IPSI.</p> <p>Berg, M. M. (2014). Modelling of natural dialogues in the context of speech-based information and control systems [Doctoral dissertation, Christian-Albrechts-Universität zu Kiel].</p> <p>Voxify, Inc. (2005). Automated Agents™ for airlines. Voxify Data Sheet (DS_Airline_May05_v04).</p> <p>Voxify, Inc. (2005). Automated Agents™ for utilities. Voxify Data Sheet (DS_Utility_May05_V03).</p> <p>Gorin, A. (2002, November 6). Semantic Information Processing of Spoken Language - How May I Help You? IEEE Communications Society North Jersey Chapter and NJIT Electrical and Computer Engineering Department.</p> <p>Goldman, E. L., Panttaja, E., Wojcikowski, A., & Braudes, R. (2001). Voice portals—Where theory meets practice. <i>International Journal of Speech Technology</i>, 4, 227–240. IJST-voice-portal</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|--------------------------|--|
| | <p>Harris, T. K., & Lewis, M. (2005). Heterogeneous multi-robot dialogues for search tools. Proceedings of the 2005 IEEE International Conference on Robotics and Automation.</p> <p>Error Handling. Internal technical documentation on speech application development and recovery strategies. Error Handling</p> <p>EDN. (2001, June 5). Speaking tour: ASR and TTS follow the money. EDN Magazine.</p> <p>Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: A tutor with dialogue in natural language. Behavior Research Methods, Instruments, & Computers, 36(2), 180–192.</p> <p>Hirschberg, J. (1999). Communication and prosody: Functional aspects of prosody. In Proceedings of the ISCA Workshop on Dialogue and Prosody.</p> <p>Jurafsky, D., & Martin, J. H. (2006). Dialogue and conversational agents. In Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition (Draft of November 12, 2006).</p> <p>Guinn, C. I. (1996). Mechanisms for mixed-initiative human-computer collaborative discourse. Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics.</p> <p>Stent, A. J., Marge, M., & Singhai, A. (2005). Evaluation methods for spoken dialog systems. IBM Research Technical Report.</p> <p>Raux, A., Langner, B., Bohus, D., Black, A. W., & Eskenazi, M. (2005). Let's Go Public! Taking a spoken dialog system to the real world. In Proceedings of Interspeech 2005. IS051938</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|--|
| | <p>Fraser, N. M. (1993). The SUNDIAL speech understanding and dialogue project: Results and implications for translation. In <i>Translating and the Computer 15: Machine Translation Today</i>. London: Aslib.</p> <p>Belvin, R., Burns, R., & Hein, C. (2001). Development of the HRL route navigation dialogue system. Proceedings of the HLT 2001 Conference.</p> <p>LumenVox LumenVox Speech Recognition Engine Documentation. LumenVox LLC. engine_help</p> <p>Hirschman, L., & Pallett, D. S. (1993). The AT&T spoken language understanding system. AT&T Technical Journal. Gupta 2004</p> <p>Fraser, N. M. (1993). Corpus-based evaluation of the Sundial system. In J. G. Neal & S. M. Walter (Eds.), <i>Natural Language Processing Systems Evaluation Workshop</i> (pp. 103–108). Rome Laboratory Technical Report RL-TR-91-362.</p> <p>Biersdorfer, J. D. (2000, August 3). Now yelling at your PC can actually be useful. <i>The New York Times</i>.</p> <p>Bohus, D., Raux, A., Harris, T. K., Eskenazi, M., & Rudnicky, A. I. (2007, April). Olympus: An open-source framework for conversational spoken language interface research. In <i>Bridging the Gap: Academic and Industrial Research in Dialog Technologies Workshop, NAACL-HLT</i>.</p> <p>Hirasawa, J., Miyazaki, N., Nakano, M., & Aikawa, K. (2000). New feature parameters for detecting misunderstandings in a spoken dialogue system. NTT Laboratories.</p> <p>Juneja, A. (2003). On the road to the creation of situation-adaptive dialogue managers. Carnegie Mellon University.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|---|
| | <p>Jurafsky, D., Shriberg, E., Fox, B., & Curl, T. (1998). Lexical, prosodic, and syntactic cues for dialog acts. In Proceedings of the ACL/COLING Workshop on Discourse and Dialogue.</p> <p>Langkilde, I. (2000). Forest-based statistical sentence generation. AT&T Labs – Research.</p> <p>Lee, K. F. (1990). Context-dependent phonetic hidden Markov models for speaker-independent continuous speech recognition. <i>IEEE Transactions on Acoustics, Speech, and Signal Processing</i>, 38(4), 599–609.</p> <p>Lemon, O., Bracy, A., Gruenstein, A., & Peters, S. (2001). The WITAS multi-modal dialogue system I. CSLI, Stanford University.</p> <p>Lemon, O., & Gruenstein, A. (2002). Multithreaded context for robust conversational interfaces: Context-sensitive speech recognition and interpretation of corrective fragments. In Proceedings of the 3rd SIGdial Workshop on Discourse and Dialogue (pp. 113–124).</p> <p>LumenVox. (n.d.). Disambiguating speech engine applications. LumenVox White Paper.</p> <p>Mast, M., Kummert, F., Ehrlich, U., Fink, G., Kuhn, T., Niemann, H., & Sagerer, G. (1994). A speech understanding and dialog system with a homogeneous linguistic knowledge base. <i>IEEE Transactions on Pattern Analysis and Machine Intelligence</i>, 16(2), 179–194.</p> <p>McTear, M. F. (2002). Spoken dialogue technology: Enabling the conversational user interface. <i>ACM Computing Surveys</i>, 34(1), 90–169.</p> <p>Möbius, B., Sproat, R., van Santen, J. P. H., & Olive, J. P. (1997). The Bell Labs German text-to-speech system: An overview. Bell Labs – Lucent Technologies.</p> <p>Nakano, M., Miyazaki, N., Hirasawa, J., Dohsaka, K., & Kawabata, T. (1999). Understanding unsegmented user utterances in real-time spoken dialogue systems. NTT Laboratories.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|--|
| | <p>Nöth, E., Batliner, A., Kießling, A., Kompe, R., & Niemann, H. (2000). VERBMOBIL: The use of prosody in the linguistic components of a speech understanding system. <i>IEEE Transactions on Speech and Audio Processing</i>, 8(5), 519–532.</p> <p>Shriberg, E., Bates, R., Stolcke, A., Taylor, P., Jurafsky, D., Ries, K., Coccaro, N., Martin, R., Meteer, M., & Van Ess-Dykema, C. (1998). Can prosody aid the automatic classification of dialog acts in conversational speech? <i>Language and Speech, Special Issue on Prosody and Conversation</i>.</p> <p>Batliner, A., Kießling, A., Kompe, R., Niemann, H., & Nöth, E. (1995). Can we tell apart intonation from prosody if we look at accents and boundaries? (Report No. 215). Friedrich-Alexander-Universität Erlangen-Nürnberg.</p> <p>Schmitz, B. (1996). A relevance-based approach to automatic interpreting (Report No. 164). Technische Universität Berlin.</p> <p>Bohus, D., & Rudnicky, A. I. (2009). The RavenClaw dialog management framework: Architecture and systems. <i>Computer Speech and Language</i>, 23(3), 332–361.</p> <p>Fetter, P. (1996). Detection and transcription of OOV words (Report No. 231). Daimler-Benz AG Research and Technology.</p> <p>Gonsalves, C. (2000, September 11). Playing host to IVAN. <i>eWeek</i>, 55.</p> <p>Du Bois, G. (2000, November 13). One Voice offers voice navigation service. <i>eWeek</i>, 16.</p> <p>Reithinger, N., & Maier, E. (1995). Utilizing statistical dialogue act processing in Verbmobil. DFKI GmbH.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|--|
| | <p>Paeschke, A., & Sendlmeier, W. F. (2000). Prosodic characteristics of emotional speech: Measurements of fundamental frequency movements. Technische Universität Berlin.</p> <p>Raux, A., Bohus, D., Langner, B., Black, A. W., & Eskenazi, M. (2006). Doing research on a deployed spoken dialogue system: One year of Let’s Go! experience. In INTERSPEECH 2006 - ICSLP (pp. 1315–1318). Carnegie Mellon University.</p> <p>Parent, G., & Eskenazi, M. (2010). Lexical entrainment of real users in the Let’s Go spoken dialog system. In INTERSPEECH 2010 (pp. 3018–3021). Carnegie Mellon University.</p> <p>Wahlster, W. (Ed.). (2006). Prosody section. In SmartKom: Foundations of multimodal dialogue systems (pp. 117–125). Springer.</p> <p>Zeißler, V., Adelhardt, J., Batliner, A., Frank, C., Nöth, E., Shi, R. P., & Niemann, H. (2006). The Prosody Module. In W. Wahlster (Ed.), SmartKom: Foundations of multimodal dialogue systems (pp. 139–150). Springer.</p> <p>Chen, F., & Jokinen, K. (Eds.). (2010). Speech technology: Theory and applications. Springer Science+Business Media.</p> <p>Freedman, R. (1999). Atlas: A plan manager for mixed-initiative, multimodal dialogue. In AAAI '99 Workshop on Mixed-Initiative Intelligence.</p> <p>Furui, S. (2010). History and development of speech recognition. In F. Chen & K. Jokinen (Eds.), Speech technology: Theory and applications (pp. 1–18). Springer.</p> <p>Rudnicky, A. I., Thayer, E., Constantinides, P., Tchou, C., Shern, R., Lenzo, K., Xu, W., & Oh, A. (1999). Creating natural dialogs in the Carnegie Mellon Communicator system. Carnegie Mellon University.</p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidation Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|--|
| | <p>Sandewall, E., Doherty, P., Lemon, O., & Peters, S. (2004). Words at the right time: Real-time dialogues with the WITAS unmanned aerial vehicle.</p> <p>Singh, S. (2000). Optimizing dialogue management with reinforcement learning: Experiments with the NJFun system.</p> <p>Smith, R. W. (1993). An empirical study of variable initiative behavior in spoken natural language dialog.</p> <p>Smith, R. W., & Hipp, D. R. (1994). Spoken natural language dialog systems: A practical approach. Oxford University Press.</p> <p>Smith, R. W., Hipp, D. R., & Biermann, A. W. (1992). An architecture for voice dialog systems based on Prolog-style theorem proving. <i>Computational Linguistics</i>, 21(3), 281–320.</p> <p>Speech Recognition Engine Spec Sheet. LumenVox.</p> <p>SpeechTEK 2005 New York Show Directory. (2005). AmComm, Inc.</p> <p>SpeechTEK 2006 New York Show Directory. (2006). AmComm, Inc.</p> <p>Wiemer-Hastings, P., Allbritton, D., & Arnott, E. (2004). RMT: A dialog-based research methods tutor with or without a head.</p> <p>Wahlster, W. (2000). Mobile speech-to-speech translation of spontaneous dialogs: An overview of the final Verbmobil system. <i>Verbmobil: Foundations of speech-to-speech translation</i>, 3-21.</p> <p>Vo, M. T., & Wood, C. (1996, May). Building an application framework for speech and pen input integration in multimodal learning interfaces. In <i>1996 IEEE International Conference on</i></p> |

VB Assets, LLC, v. Samsung Electronics Co., LTD.
No. 2:24-cv-00828-JRG (E.D. Tex.)
Defendants’ Invalidity Contentions – Appendix G-3

| Challenged Claims | Other Bases for Why the Challenged Claims are Patent-Ineligible |
|-------------------|--|
| | <p><i>Acoustics, Speech, and Signal Processing Conference Proceedings</i> (Vol. 6, pp. 3545-3548). IEEE.</p> <p>Aust, H., Oerder, M., Seide, F., & Steinbiss, V. (1995). The Philips automatic train timetable information system. <i>Speech Communication</i>, 17(3-4), 249-262.</p> <p>Smith, R. W., & Hipp, D. R. (1994). <i>Spoken natural language dialog systems: A practical approach</i>. Oxford University Press.</p> <p>Wahlster, W. (Ed.). (2000). <i>Verbmobil: Foundations of Speech-to-Speech Translation</i>. Springer Science & Business Media.</p> <p>Wahlster, W. (2001). Pervasive Speech and Language Technology. In <i>Informatics: 10 Years Back, 10 Years Ahead</i> (pp. 274-293). Berlin, Heidelberg: Springer Berlin Heidelberg.</p> <p>Whittaker, S., & Stenton, P. (1989, April). User studies and the design of natural language systems. In <i>Fourth Conference of the European Chapter of the Association for Computational Linguistics</i>.</p> <p>Wu, L., Oviatt, S. L., & Cohen, P. R. (1999). Multimodal integration-a statistical view. <i>IEEE Transactions on Multimedia</i>, 1(4), 334-341.</p> |