

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.,
SAMSUNG ELECTRONICS, AMERICA, INC., and GOOGLE LLC

Petitioners

v.

CERENCE OPERATING COMPANY,

Patent Owner

Case IPR2024-00978

U.S. Patent No. 7,395,078

**DECLARATION OF MARK R. LANNING IN SUPPORT
OF PETITION FOR INTER PARTES REVIEW
OF U.S. PATENT NO. 7,395,078**

TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION	1
II. BACKGROUND AND QUALIFICATIONS	1
III. MATERIALS REVIEWED	6
IV. LEVEL OF ORDINARY SKILL IN THE ART	8
V. RELEVANT LEGAL STANDARDS	9
VI. SUMMARY OF OPINIONS	13
VII. OVERVIEW OF THE '078 PATENT	13
VIII. OVERVIEW OF THE PRIOR ART	17
A. Analogous Art	17
B. <i>Dolan</i>	18
C. <i>Freedman</i>	34
D. <i>Halonen</i>	37
IX. CLAIM CONSTRUCTION	42
X. SPECIFIC GROUNDS FOR CHALLENGE	42
A. Ground 1A: <i>Dolan</i> Renders Obvious Claims 1-3 and 8-11	43
1. Claim 1	43
a. 1[pre]: “A method of sending a voice message via a mobile communication device, said method comprising:”	43
b. 1[a]: “receiving an utterance from a user of the mobile communication device;”	47
c. 1[b]: “generating a non-text representation of the received utterance;”	48
d. 1[c]: “inserting at least some of the non-text representation into a body of a text message; and”	51
e. 1[d]: “sending the text message including the at least some of the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device.”	54

2.	Claim 2: “The method of claim 1, wherein the mobile communication device is a cellular phone.”	65
3.	Claim 3: “The method of claim 1, wherein inserting at least some of the non-text representation involves inserting all of the non-text representation into the body of the text message.” ...	67
4.	Claim 8: “The method of claim 1, wherein the wireless messaging channel is an SMS channel and the text message is an SMS message.”	68
5.	Claim 9: “The method of claim 1, further comprising including an indicator with the text message identifying the text message as containing a non-text representation of the utterance.”	68
6.	Claim 10: “The method of claim 1, wherein the non-text representation is a compressed version of the received utterance.”	70
7.	Claim 11	71
a.	11[a]: “The method of claim 1, further comprising dividing the non-text representation into sections and inserting each section into a body of a different corresponding text message of a plurality of text messages, wherein the first-mentioned text message is one of said plurality of text messages and”	71
b.	11[b]: “further comprising sending the plurality of text messages over the wireless messaging channel from the mobile communication device to the recipient’s device.”	74
B.	Ground 1B: <i>Dolan</i> in view of <i>Freedman</i> Renders Obvious Claims 1, 4-7.....	75
1.	Claim 1	76
a.	1[pre] & 1[a]: 1[pre] (“A method of sending a voice message via a mobile communication device, said method comprising:”), and 1[a] (“receiving an utterance from a user of the mobile communication device;”).....	76
b.	1[b]: “generating a non-text representation of the received utterance;”	76

c.	1[c] & 1[d]: 1[c] (“inserting at least some of the non-text representation into a body of a text message; and”), and 1[d] (“sending the text message including the at least some of the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device.”).....	84
2.	Claim 4: “The method of claim 1, wherein generating the non-text representation of the received utterance involves performing recognition on a signal derived from the received utterance to generate a string of symbols, wherein the string of symbols is the non-text representation.”	85
3.	Claim 5: “The method of claim 4, wherein the symbols in the string of symbols are selected from the group consisting of phonemes, diphones, triphones, syllables, and demisyllables.”	85
4.	Claim 6: The method of claim 4, wherein the recognition is phoneme recognition and the symbols in the generated string of symbols are phonemes.	87
5.	Claim 7: “The method of claim 1, wherein generating the non-text representation of the received utterance involves performing phoneme recognition on a signal derived from the received utterance to generate a phoneme string, wherein the phoneme string is the non-text representation.”	87
C.	Ground 2A: <i>Dolan</i> in view of <i>Halonen</i> Renders Obvious Claims 12, 15-17.....	87
1.	Claim 12	87
a.	12[pre]: “A mobile communication device for sending a voice message, said mobile communication device comprising:”	87
b.	12[a] through 12[d]: 12[a] (“a processor system;”), 12[b] (“a microphone for receiving an utterance from a user of the mobile communication device;”), 12[c] (“a transceiver; and”), 12[d] (“memory storing code which when executed on the processor system causes the mobile communication device to:”)	88

c.	12[e]: “generate a non-text representation of the received utterance;”	96
d.	12[f]: “insert the non-text representation into a body of a text message; and”	97
e.	12[g]: “send the text message including the non-text representation via the transceiver over a wireless messaging channel from the mobile communication device to a recipient’s device.”	97
2.	Claim 15: “The mobile communication device of claim 12, wherein the wireless messaging channel is an SMS channel and the text message is an SMS message.”	100
3.	Claim 16: “The mobile communication device of claim 12, wherein the non-text representation is a compressed version of the received utterance.”	100
4.	Claim 17	100
a.	Claim 17[a]: “The mobile communication device of claim 12, wherein the code when executed on the processor further causes the mobile communication device to: divide the non-text representation into sections; insert each section into a body of a different corresponding text message of a plurality of text messages, wherein the first-mentioned text message is one of said plurality of text messages; and”	100
b.	17[b]: “send the plurality of text messages over the wireless messaging channel from the mobile communication device to the recipient’s device.”	101
D.	Ground 2B: <i>Dolan</i> in view of <i>Halonen</i> and <i>Freedman</i> Renders Obvious Claims 12, 15-17	102
1.	Claim 12	102
a.	12[pre] through 12[d]: 12[pre] (“A mobile communication device for sending a voice message, said mobile communication device comprising:”), 12[a] (“a processor system;”), 12[b] (“a microphone for receiving an utterance from a user of the mobile communication device;”), 12[c] (“a transceiver; and”), 12[d] (“memory storing code	

	which when executed on the processor system causes the mobile communication device to:”).....	102
b.	12[e]: “generate a non-text representation of the received utterance;”	103
c.	12[f] & 12[g]: 12[f] (“insert the non-text representation into a body of a text message; and”), and 12[g] (“send the text message including the non-text representation via the transceiver over a wireless messaging channel from the mobile communication device to a recipient’s device.”)	104
2.	Claim 13: “The mobile communication device of claim 12, wherein generating the non-text representation of the received utterance involves performing recognition on a signal derived from the received utterance to generate a string of symbols, the string of symbols being the non-text representation.”	105
3.	Claim 14: “The mobile communication device of claim 13, wherein the symbols in the string of symbols are selected from the group consisting of phonemes, diphones, triphones, syllables, and demisyllables.”	106
XI.	SECONDARY CONSIDERATIONS	106
XII.	CONCLUSION	107

I, Mark R. Lanning, declare that:

I. INTRODUCTION

1. I have been retained by the Samsung Electronics Co., Ltd. and Samsung Electronics America, Inc. (“Samsung” or “Petitioners”) as an independent expert consultant in this proceeding before the United States Patent and Trademark Office (“PTO”).

2. I am being compensated at a rate of \$550/hour for my services in this proceeding, which is my regular and customary rate.

3. My compensation is in no way contingent on the nature of my finding, the presentation of my findings in testimony or this declaration, or the outcome of this or any other proceeding. I have no other interest in this proceeding.

4. I have been asked to consider whether certain references disclose and/or suggest the features recited in claims 1-17 of U.S. Patent No. 7,395,078 (“the ’078 patent”) (Ex. 1001).¹ My opinions are set forth below.

II. BACKGROUND AND QUALIFICATIONS

5. I am an independent consultant. All of my opinions stated in this declaration are based on my own personal knowledge and professional

¹ Where appropriate, I refer to exhibits which I understand will be attached to the petition for *inter partes* review of the ’078 patent (the “Petition”).

judgment. In forming my opinions, I have relied on my knowledge and extensive experience in the field of telecommunications including development of software and hardware for implementation of circuit switched networks, multiple generations of cellular networks (including the 2G and 3G cellular networks), and packet switched networks.

6. I am over 18 years of age and, if I am called upon to do so, I would be competent to testify as to the matters set forth herein. A copy of my current curriculum vitae, which details my education and professional experience, is included as Ex. 1003 in this proceeding. The following provides an overview of some of my experience that is relevant to the matters set forth in this declaration.

7. I have extensive experience with 2G, 3G and 4G LTE cellular networks and their standards which provides me with an understanding of the subject matter described and claimed in the '078 patent. Further, I generally have extensive experience in the field of telecommunications including circuit-switched networks, multiple generations of cellular networks, and packet-switched networks.

8. I received a B.S. in Computer Science from Southern Methodist University ("SMU") in 1983.

9. I am currently the owner and president of two consulting companies: Telecom Architects, Inc. and Reticle Consulting, LLC. Telecom

Architects (created in 1999) provides consulting services to fixed and wireless telecom service providers and their equipment suppliers. Reticle Consulting (created in 2009) provides specialized consulting and expert services for forensic software analysis and software source code comparison for software misappropriation cases.

10. My relevant experience, which spans over 40 years, includes work as an architect of various telecommunications systems and work developing equipment used in telecommunications systems. This experience includes extensive design, implementation, and testing work on the wireless interface functionality (between the base station and mobile phones) for multiple generations of cellular standards.

11. I began my professional career in 1983 at DSC Communications Corp. (“DSC”), now a part of Alcatel, as a software development manager on the team responsible for converting DSC’s PSTN telephone switch into a Mobile Switching Center (“MSC”) for Motorola to sell as a part of their cellular product offering in the U.S. and many other countries. In 1985, I was hired by Telinq, as its director of software development and later was promoted to vice president of hardware and software development. Telinq was a startup company building high-speed transmission multiplexer products.

12. In 1987, I was hired by the Telecom Division of Tandem Computers, Inc. (now a part of HP) as its director of hardware and software development. Tandem Telecom built products for phone companies and wireless carriers to process many different types of customer transactions including 800 number and Home Location Register (HLR) applications.

13. In 1991, I left Tandem Telecom and I created my first consulting company, I.N. Solutions, Inc.,² to provide consulting services for telecom equipment suppliers and network operators. For one of my first contracts, I worked as a hardware and software consultant to Motorola for its “SuperCell” base station product and as a consultant to British Telecom to upgrade its current analog cellular network, and I was one of the network architects responsible for the design and rollout of British Telecom’s Global System for Mobile Communications (“GSM”) network known as Cellnet.

14. I had a subsequent consulting contract with British Telecom’s Cellnet division that began in the early 1990s and lasted over seven years. As a program manager, I was responsible for the implementation of the Short Message Service (“SMS”) on Cellnet’s new GSM network. This role included

² I.N. Solutions, Inc. is no longer active. I worked contracts with I.N. Solutions, Inc. until forming Telecom Architects, Inc. in 1999.

me personally working with suppliers of Cellnet's Short Message Service Center ("SMSC"), Aldiscon; Base Station Systems, Nokia and Motorola; Mobile Switching Centers, Ericsson; and cellular phones, Nokia and Ericsson, to define and roll out mobile terminated and mobile originated SMS functionality.

15. Since 1995, I have also consulted with other equipment suppliers and network operators for second generation (2G) and third generation (3G) Code Division Multiple Access ("CDMA") network architecture and equipment design and implementation, providing consulting services to companies such as Sprint, Nextel, Nokia, and Northern Telecom Ltd. Specifically, while consulting to Nextel, that later became part of Sprint and is now owned by T-Mobile US, I was one of the network architects for Nextel's iDEN network, and its evolution from 2G to 3G (UMTS and CDMA2000) and for which one of my further responsibilities was to define the network and mobile phone functionality required to support Nextel's Multimedia Messaging Service ("MMS") and advanced data communications capability associated with its push-to-talk phones.

16. In addition to my cellular experience listed above, for at least the past ten years, I have kept up to date my knowledge of the present-day 3G, 4G, and 5G cellular networks and wireless standards and their associated equipment and protocols. I have done this through my study of each new release of these

standards, technical books, and trade publications as well as my expert work in legal cases, which has involved evaluating the functionality of many different types of network equipment, mobile devices, and baseband chipsets. I have also evaluated thousands of cellular technology patents and most of these were for cellular phones and/or their base stations.

17. I am a member of the Institute of Electrical and Electronics Engineers (“IEEE”), including the IEEE Standards Association. I am also a member of the Association for Computing Machinery (“ACM”). While employed at DSC, I was a member of the ANSI T1 and T1X1 standard groups responsible for the definition and standardization of the Advanced Intelligent Network (“AIN”) and Signaling System 7 (“SS7”) protocol, which is the base protocol used to carry SMS messages in a cellular network.

III. MATERIALS REVIEWED

18. The opinions contained in this declaration are based on the documents I reviewed, my professional judgment, as well as my education, experience, and knowledge regarding implementation of messaging architecture and protocols within cellular networks.

19. In forming my opinions expressed in this declaration, I reviewed the following materials:

- U.S. Patent 7,395,078 (“’078*Pat.*”) (Ex. 1001);

- File History of U.S. Patent 7,395,078 (Ex. 1004);
- Canadian Patent CA2275828C (“*Dolan*”) (Ex. 1005);
- U.S. Publication 2004/0073423A1 (“*Freedman*”) (Ex. 1006);
- U.S. Patent 5,887,254A (“*Halonen*”) (Ex. 1007);
- Excerpt from The New Oxford American Dictionary (2001)

(Ex. 1008);

- U.S. Patent 5,870,427 (“*Tiedemann*”) (Ex. 1009);
- U.S. Patent 6,081,724 (“*Wilson*”) (Ex. 1010);
- U.S. Patent 5,963,548 (“*Virtanen*”) (Ex. 1011);
- U.S. Patent 5,953,322 (“*Kimball*”) (Ex. 1012)
- Excerpts from Computer Dictionary (4th ed. 1999) (Ex. 1013);

and any other materials I refer to in this declaration in support of my opinions.

20. All of the opinions contained in this declaration are based on the documents I reviewed and my knowledge and professional judgment. My opinions have also been guided by my appreciation of how a person of ordinary skill in the art would have understood the claims and the specification of the ’078 patent at the time of the alleged invention, which I have been asked to initially consider as early as April 20, 2004. My opinions reflect how one of ordinary skill in the art would have understood the ’078 patent, the prior art to the patent, and the state of the art at the time of the alleged invention.

21. It is my opinion that certain references disclose and/or suggest all the features recited in claims 1-17 (“Challenged Claims”) of the ’078 patent, as I discuss in detail below.

IV. LEVEL OF ORDINARY SKILL IN THE ART

22. I have been informed and understand that, in the context of an invalidity analysis, a person having ordinary skill in the art is a hypothetical person who looks to prior art at the time of the invention. I further understand that the factors that may be considered in determining the level of ordinary skill include: (1) the problems encountered in the art; (2) the prior art solutions to the problems encountered in the art; (3) the rapidity of innovation; (4) the sophistication of the technology; and (5) the education level of active workers in the field. I understand that these factors need not all be considered for the analysis and that one or more of these factors may control.

23. I was asked to provide my opinion on the level of one of ordinary skill in the art with respect to the alleged invention of the ’078 patent in April 2004. Based on my consideration of the factors above, I believe a person of ordinary skill in the art would have had would have had a Bachelor’s of Science Degree (or equivalent) in electrical engineering, computer science, or computer engineering, or a related technical field and two years of experience working with mobile or wireless communications networks such as cellular networks, or

a related technical field, and familiarity with short message service protocols. A greater amount of education could compensate for fewer years of work experience, and vice versa.

24. As of April 20, 2004, I met, and in fact exceeded, the qualifications of a person of ordinary skill in the art. To be clear, all of my opinions in this declaration are from the perspective of one of ordinary skill in the art as I have defined it here during the relevant timeframe.

V. RELEVANT LEGAL STANDARDS

25. I am not an attorney and offer no legal opinions, but in the course of my work, I have had experience studying and analyzing patents and patent claims from the perspective of a person skilled in the art.

26. For the purposes of this declaration, I have been informed about certain aspects of the law that are relevant to forming my opinions. My understanding of the law is as follows.

27. Petitioner's counsel has informed me that for the prior art to inherently disclose a claimed limitation, the prior art need not expressly disclose the limitation, so long as the claimed limitation necessarily flows from a disclosure in the prior art. I also understand that it is acceptable to examine evidence outside the prior art reference (extrinsic evidence) in determining

whether a feature, while not expressly discussed in the reference, is necessarily present in that reference.

28. Petitioner's counsel has informed me that a patent claim can be considered to have been obvious to a person having ordinary skill in the art at the time the application was filed. I am informed that this means that, even if all of the requirements of a claim are not found in a single prior art reference, the claim is not patentable if the differences between the subject matter in the prior art and the subject matter in the claim would have been obvious to a person having ordinary skill in the art at the time of the invention.

29. I have been informed by Petitioners' counsel that a determination of whether a claim would have been obvious should be based upon several factors, including, among others:

- the level of ordinary skill in the art at the time the application was filed;
- the scope and content of the prior art;
- what differences, if any, existed between the claimed invention and the prior art; and
- any "secondary indicia" of non-obviousness, if they are of record.

30. I have been informed by Petitioners' counsel that a single reference can render a patent claim obvious if any differences between that reference and

the claims would have been obvious to a person having ordinary skill in the art. Alternatively, I understand that the teachings of two or more references may be combined in the same way as disclosed in the claims, if such a combination would have been obvious to one having ordinary skill in the art. In determining whether a combination based on either a single reference or multiple references would have been obvious, the following are examples of approaches and rationales that may be considered:

- whether the teachings of the prior art references disclose known concepts combined in familiar ways, and when combined, would yield predictable results;
- whether a person of ordinary skill in the art could implement a predictable variation, and would see the benefit of doing so;
- whether the claimed elements represent one of a limited number of known design choices, and would have a reasonable expectation of success by those skilled in the art;
- whether a person of ordinary skill would have recognized a reason to combine known elements in the manner described in the claim;
- whether there is some teaching or suggestion in the prior art to make the modification or combination of elements claimed in the patent; and

- whether the innovation applies a known technique that had been used to improve a similar device or method in a similar way.

31. I understand that “secondary indicia” of non-obviousness may include certain objective factors, such as: commercial success of products practicing the claimed invention; long-felt but unsolved need; teaching away; unexpected results; copying; and praise by others in the field. These factors are generally referred to as “secondary considerations” or “objective indicia” of non-obviousness. I understand, however, that for such objective evidence to be relevant to the (non)obviousness of a claim, there must be a causal relationship (called a “nexus”) between the claim and the evidence and that this nexus must be based on a novel element of the claim rather than something in the prior art. I also understand that even when they are present, secondary considerations may be unable to overcome primary evidence of obviousness (such as motivation to combine with predictable results) that is sufficiently strong.

32. I have been informed by Petitioners’ counsel that one of ordinary skill in the art has ordinary creativity and is not an automaton. Petitioners’ counsel has also informed me that in considering obviousness, obviousness may not be determined using the benefit of hindsight, including hindsight derived from the patent being considered.

33. Petitioners' counsel has also explained to me what a Markush group is in a claim. My understanding is that a Markush group is written in the form "a group consisting of [A, B, and C]," and that for purposes of infringement, an alleged infringer need to practice only one of the members of the group to infringe. Additionally, it was explained to me that to invalidate a claim containing a Markush group, the prior art need only show one of the members of the group.

VI. SUMMARY OF OPINIONS

34. For the reasons I discuss below, it is my opinion that claims 1-17 of the '078 patent are rendered obvious by the prior art.

VII. OVERVIEW OF THE '078 PATENT

35. The '078 patent, titled "Voice over short message service," relates to "[a] method of sending a voice message via a mobile communication device" comprising "receiving an utterance from a user of the mobile communication device," "generating a non-text representation of the received utterance," "inserting the non-text representation into a body of a text message," and "sending the text message over a wireless messaging channel from the mobile communication device to a recipient's device." Ex. 1001, Abstract. The method involves "sending and receiving spoken or audio information over the SMS

network available in cellular phones.” *Id.*, 3:33-35; *see also id.*, 2:22 (identifying the “mobile communicate device is a cellular phone”), 2:48 (same).

36. SMS, which stands for Short Message Service, is a text messaging service, and SMS messages were and continue to be a common form of text messaging.

37. The ’078 patent describes some of the benefits of using SMS networks, including avoidance of high costs, market demand and efficiency of being able to text while on a voice call:

In response to the high costs of voice calls, text based mobile-to-mobile messaging called SMS, or Short Message Service, has become heavily used in some markets, particularly amongst younger demographics. SMS enables a user to transmit and receive short text messages at any time, regardless of whether a voice call is in progress. The user typically types in the message text through the small keyboard that is provided on the device. The messages are hardware limited to 160 characters, and are sent as packets through a low bandwidth, out-of-band message transfer channel. This allows for facile communication with minimal burden on the wireless network.

Id., 1:37-49.

38. The ’078 patent also contemplates the benefit of using voice messaging for those demographics having a low literacy rate:

But in emerging markets, such as India which has a very large population, the adult literacy rate is roughly 60% and thus a large number of people are not sufficiently literate to type text messages into the cell phone. Thus, for many consumers in such markets who can neither compose nor read a message, SMS text-messaging as a communication mode is not an effective alternative. At least some of the embodiments described herein provide a mechanism by which such consumers can nevertheless use the lower-cost, non-voice wireless communication channels for verbal communications instead of text messaging.

Id., 2:2-13.

39. The '078 patent further describes the use of legacy wireless network systems like GSM, TDMA, and CDMA to utilize SMS messaging over a text/data channel and indicates the significant source of SMS revenue:

Most legacy wireless network systems such as GSM, TDMA, and CDMA have a text/data channel capable of sending and receiving SMS, so the infrastructure for this service already exists even in emerging markets in developing countries. Some estimates now place the global number of SMS messages at nearly 40 billion messages per month. It is thought that SMS is now the most significant source of non-voice based revenue to wireless network operators worldwide. As a result carriers are very interested in promoting the use of SMS. Indeed, network operators in developing markets may limit the implementation of more advanced voice network infrastructures due to the large revenues associated with text messaging.

Id., 1:50-61.

40. Moreover, the '078 patent contemplates the use of multiple different algorithms to generate a compressed audio message to be sent by message. One method of compression can use a “phonetic recognition algorithm in the phone” to generate a non-text representation of a an utterance spoken by the user into a cellular phone. *Id.*, 3:35-37. Alternatively, “other algorithms could be used which simply perform compression without performing any recognition.” *Id.*, 6:56-58. For example, “instead of using a phoneme recognizer, a vocoder could be used to generate the compressed representation and then that compressed representation would be inserted into the body of the SMS message. In other words, ***any algorithm that produces a non-text representation suitable for sending over SMS or another non-voice channel could be employed.***” *Id.*, 6:58-64.³ In view of these, the '078 patent discloses inserting into an SMS message a “stored phoneme string” or a vocoder-generated “compressed representation.” *Id.*, 4:55-59, 6:58-64 (same). The '078 patent also describes programming an SMS application “to generate a sequence of multiple SMS messages to handle longer utterances for which the non-text representation will not fit into the body of a single message.” *Id.*, 6:15-18.

³ I have added all emphases/annotations/color unless I note otherwise.

41. The '078 patent also describes that on the receiving end, an SMS application in the cellular phone “processes the message for presentation to the user.” *Id.*, 5:59-62. “A flag within the received SMS message identifies the contents of the SMS message” as a non-text representation that is “processed by the synthesizer to generate an audio signal.” *Id.*, 5:62-65. The non-text representation is passed “to a synthesizer 220 within the cell phone, along with any stored parameters which were supplied to control synthesizer 220 and the way it generates the sounds.” *Id.*, 6:2-5. “The synthesizer plays through the cell phone speaker the sequence of sounds” that represent the non-text representation. *Id.*, 6:10-12. “The SMS application on the recipient’s side of the connection would also need to be programmed to recognize that the received non-text representation is to be constructed by concatenating the contents of more than one SMS message.” *Id.*, 6:24-28.

42. The claims of the '078 patent are directed, generally speaking, to methods and devices for sending, over a wireless messaging channel, a non-text representation of a user’s utterance in a body of a text message.

VIII. OVERVIEW OF THE PRIOR ART

A. Analogous Art

43. I have analyzed the prior art references used in the grounds below, and it is my opinion that they are analogous to the '078 patent because they are

within the same field of endeavor and reasonably pertinent to at least one problem addressed by the '078 patent.

44. In particular, *Dolan* (like the '078 patent) is directed to methods and techniques for compression of audio information including a user's voice. *Dolan* further discusses doing this by using mobile phones or similar devices that work over a wireless communication channel using communication protocols such as SMS to send compressed audio information.

45. *Freedman*, while similar to *Dolan* in many ways, differs in that *Freedman* sends user speech by a phoneme-based technique that compresses the audio creating one type of non-text representations that is discussed in the '078 patent.

46. As for *Halonen*, it provides those components of mobile phones that allows for the implementation of the methods as used in the systems disclosed in the '078 patent, *Dolan* and *Freedman*.

B. *Dolan*

47. *Dolan* is a Canadian patent titled "Using Discrete Message-Oriented Services to Deliver Short Audio Communications." Ex. 1005, (54).



Office de la Propriété
Intellectuelle
du Canada
Un organisme
d'Industrie Canada

Canadian
Intellectual Property
Office
An agency of
Industry Canada

CA 2275828 C 2003/01/14

(11)(21) **2 275 828**

(12) **BREVET CANADIEN
CANADIAN PATENT**

(13) **C**

(22) Date de dépôt/Filing Date: 1999/06/21

(41) Mise à la disp. pub./Open to Public Insp.: 2000/02/19

(45) Date de délivrance/Issue Date: 2003/01/14

(30) Priorité/Priority: 1998/08/19 (09/136,678) US

(51) Cl.Int.⁶/Int.Cl.⁶ H04M 3/50, H04Q 7/00

(72) Inventeur/Inventor:
DOLAN, MICHAEL FRANCIS, US

(73) Propriétaire/Owner:
LUCENT TECHNOLOGIES, INC., US

(74) Agent: KIRBY EADES GALE BAKER

(54) Titre : UTILISATION DE SERVICES ORIENTES MESSAGES DISCRETS POUR L'ETABLISSEMENT DE COURTES COMMUNICATIONS AUDIO

(54) Title: **USING DISCRETE MESSAGE-ORIENTED SERVICES TO DELIVER SHORT AUDIO COMMUNICATIONS**

Ex. 1005, Cover (annotated).

48. *Dolan* is directed to a method and system for delivery of short audio communications (such as a spoken phrase) to a remote end-user devices using discrete message-oriented services like the well-known SMS. More particularly, *Dolan* discloses that the input audio information is first processed by an audio encoding and compression function, then the compressed audio data undergoes a Short Message creation function that creates a Short Message. Once the Short Message is created, it can be transmitted to one or more remote end-user devices, e.g., a mobile telephone.

49. *Dolan* discloses compressing captured audio information (such as a spoken phrase).

In a preferred embodiment of the invention, a segment of audio information is first captured, *the captured information is encoded by means of any audio compression technology known in the art, such as the technology used in voice coding schemes on wireless networks*, and the compressed information is stored at a message center or other facility that provides mailbox services to end users. *Many techniques for compressing and encoding voice signals for digital transmission are well-known in the art, and it is contemplated that any of these techniques may be used to perform the audio compression function of the invention.*

Id., 7:9-16.

50. *Dolan* discloses sending the compressed audio communications to a remote end-user device using a discrete message-oriented service, such as SMS:

In one embodiment of the invention, audio information is captured, the captured information is encoded by means of audio compression technology, and the encoded audio information is stored at a message center or other facility that provides mailbox services to end users. Alternatively, the received audio information may be stored at the message center in an uncompressed format and compressed at a later time, such as at the time of delivery. *The compressed audio information is transmitted by means of Short Message Service or another discrete message-oriented service to the end user. The Short Message containing the compressed audio information is received at the end user device*, the information is decompressed, and the decompressed audio information is played back to the

intended message recipient. Message delivery can be accomplished in parallel with some other ongoing activity at the end user device.

Id., 3:9-19; *see also id.*, Abstract (“Delivery of short audio communications to remote end user devices uses discrete message-oriented service messages for delivery of the audio messages. . . . The compressed audio information is transmitted by means of Short Message Service or another discrete message-oriented service to the end user device, where it is decompressed and played to the recipient.”).

In a preferred embodiment of the system, the audio information is delivered to a Message Center Audio Compression Function within a Message Center, where it is encoded and compressed digitally. ***The compressed audio information is formatted into one or more Short Messages*** by the Message Center Short Message Creation Function. ***Once one or more Short Messages are created to contain the compressed audio information, they are stored by a Message Center Short Message Storage Function until the End User Device is available for delivery.*** In a preferred embodiment, the Message Center Short Message Storage Function is implemented directly by a computing device, but other technologies known in the art are suitable.

Id., 4:18-26.

The compressed audio information is next transmitted by means of Short Message Service or another discrete message-oriented service to the end user. The Short Message containing the compressed audio information is received at an end user device,

the audio information is decompressed, and the decompressed audio information is played back to the end user (the intended message recipient). *The transmission scheme used by the preferred embodiment of this invention is Short Message Service or any other discrete message-oriented service* wherein Short Messages are captured on behalf of an end user, stored in a messaging center until the end user device is available to receive delivery, and then delivered digitally to the end user, such as the user of a digital wireless telephone device. Commonly, such delivery is done in a connectionless manner while the end user device is not otherwise in use by the user, but it can also be accomplished in parallel with, and independently of, some other ongoing activity at the end user device.

Id., 7:26-8:10.

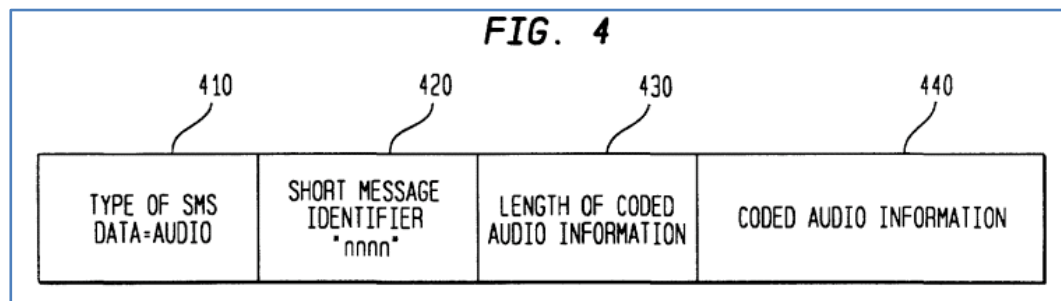
51. *Dolan* discloses delivering the Short Message(s) using various export techniques for end-to-end delivery of digital data, including wireless:

As shown in Fig 1, in a preferred embodiment of the system the audio information 102, such as a spoken phrase, is delivered to the Message Center Audio Compression Function 110 within a Message Center 106, where it is encoded and compressed digitally. The Message Center 106 will commonly be some form of voice mailbox system that provides messaging services and/or paging to the user and *is attached to the intervening network by any of the variety of export techniques known in the art for providing end-to-end delivery of digital data, including, but not limited to: Internet, signaling systems, and wireless (radio, etc.).*

Id., 10:27-11:5.

52. *Dolan* discloses that the end user device to which the Short Message is delivered “is conceived as a digital mobile telephone handset.” *Id.*, 4:12-13.

53. Figure 4 below shows the major message fields that *Dolan* describes as comprising a Short Message.



Ex. 1005, FIG. 4.

Dolan provides the following descriptions with respect to the above fields in figure 4:

The Short Message containing all or portions of the encoded audio information also contains a type field that indicates the contents of the Short Messages, as well as the type of audio decompression technique necessary to reconstitute the audio information. As seen in Fig. 4, the major fields in a Short Message Service message used to carry coded audio information are a ***Type of Data field 410***, indicating that the associated message is an Audio message and including information on the type of translator needed to decompress and decode the message field in order to retrieve the audio message, a ***Short Message Identifier field ("nnnn") 420*** that is specified by the SMS or other standard, a

Length field 430 indicating the overall coded message length, and the *coded audio information 440*. This embodiment of a message according to the invention assumes that all of *the coded audio information 440, plus the Type of Data 410, Identifier 420, and Length 430 fields, will fit into a single SMS message* sent across one or more intervening networks between a Message Center and the end user's equipment (e.g., a mobile phone, PC, etc.).

Id., 8:11-24.

54. In those instances where audio information does not fit in a single text message (e.g., “specified by the SMS standard according to the existing SMS conventions,” *id.*, 8:25-9:5), *Dolan* teaches that its system will segment and reassemble those audio messages allowing for complete and coherent delivery to the recipient. For example, *Dolan* discloses:

It is possible for the audio information received to be too large to be contained in a single Short Message, even though the end user device and the message center might be capable of handling larger messages. In such a case, segmentation and reassembly techniques commonly known in the art are applied to store and/or deliver the audio information to the end user device. The handling of a Short Message containing information too long to be transmitted in a single message across the intervening network(s) between the Message Center and the end user's equipment can be done in at least two ways. In one, the Short Message Service provides segmentation and reassembly of any Short Message that exceeds intervening network size limits.

Employment of this well known technique requires no special features for handling coded audio information. *In a second, the segmentation and reassembly process is carried out by the two endpoints which are transmitting and receiving the coded audio information. In this method, well known segmentation and reassembly techniques can also be applied.*

Id., 9:16-29.

If the audio information is too large to be contained in a single Short Message, segmentation and reassembly techniques commonly known in the art are applied to store and/or deliver the audio information to the end user device. Coded audio information that is too long for a single Short Message is broken into segments, each of which is placed into a separate Short Message having the fields: Type of Data; Short Message Identifier; Total Segments; Number of Enclosed Segment; Segment Length; and Coded Audio Information Segment. The Total Segments and Number of Enclosed Segment fields provide the means to identify and reassemble the segments of the coded audio information into the correct order at the end user device. The Length field contains the length of the associated Segment of coded audio information, and the Coded Audio Information field contains the segment itself.

Id., 3:26-4:9.

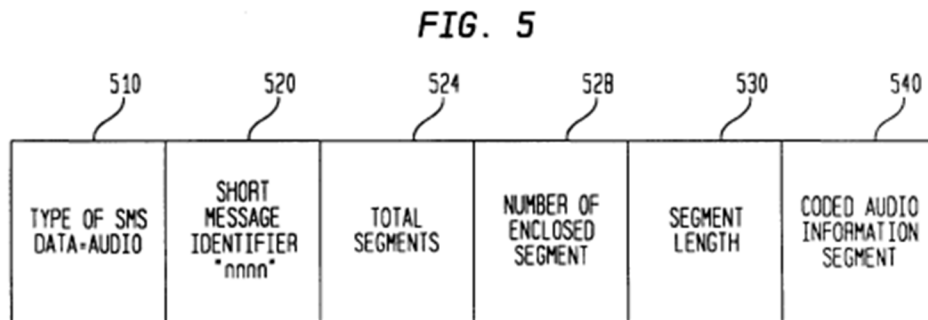
55. *Dolan* further explains that SMS has this segmentation/reassembly built in, but alternatively, the segmentation/reassembly can be included by

Dolan:

As previously described, if the compressed audio information is too long to be contained in a single Short Message, it can be segmented into multiple Short Messages utilizing any common methodology known in the art. *For example, SMS has a segmentation and reassembly feature already built in. Alternatively, a separate segmentation/reassembly capability can be provided as part of the invention, operating to divide the audio message into segments and put the segments into separate Short Messages*, each Short Message having an indicator signaling that the segment contained within the Short Message is a particular segment number out of some total number of segments. . . .

Id., 11:22-12:2.

56. Figure 5 below describes the major fields of Short Messages that require segmentation and reassembly:



Ex. 1005, FIG. 5.

Dolan provides the following descriptions with respect to the above fields in figure 5:

As shown in Fig. 5, coded audio information that is too long for a single Short Message is broken into segments, for example, into Segment 1 of 3, Segment 2 of 3, and Segment 3 of 3. Each of these segments is then placed into a separate Short Message having the following fields: ***Type of Data= Audio (again including identification of the type of translator necessary to reconstitute the audio information) 510; Short Message Identifier "nnnn" 520; Total Segments 524; Number of Enclosed Segment 528; Segment Length 530; and the Coded Audio Information Segment 540.*** The Type of Data field 510 is the same as for a non-segmented message. The Short Message Identifier 520 for a segmented message provides both a means of identifying the particular activity between the Message Center and end user, and a means of identifying the messages containing the various segments of the same coded audio information message. ***The Total Segments field 524 and the Number of Enclosed Segment field 528 provide the means to identify and then correctly reassemble the segments of the coded audio information into the correct order at the end user device.*** The Length field 530 contains the length of the associated segment of coded audio information, and the Coded Audio Information field 540 contains the segment itself.

Id., 10:1-16.

57. While *Dolan* in the first instance discusses the functionalities of capturing an audio message, compressing the audio message, inserting the audio message into one or more Short Messages, and transmitting the Short Message(s) to an end user device with reference to a "Message Center," I note that *Dolan* also discloses that "[t]he mobile user may also be the initiator of a

short audio message utilizing this invention, either as a reply to a received message or independently.” *Id.*, 15:4-5; *see also id.*, 6:27-7:1 (“[The invention] can be built into any message center ***or mobile phone*** that supports digital voice encoding and discrete message-oriented service.”). Indeed, *Dolan* discloses an embodiment with reference to figure 3 in which an end user device (e.g., a digital mobile telephone handset) receives and compresses audio information, inserts the compressed audio information into one or more Short Messages (e.g., SMS message(s)), and transmits the Short Message to another end user device. Specifically, *Dolan* describes figure 3 and the flow of audio information that travels from the End User Device 120 to the Message Center 106 as follows:

The mobile user may also be the initiator of a short audio message utilizing this invention, either as a reply to a received message or independently. ***As illustrated in Fig. 3, the flow of the audio information is from the End User Device 120 to the Message Center 106 for subsequent delivery to a second end user.*** That subsequent delivery may be accomplished by reconstitution of the compressed audio information and optional storage in a voice mailbox, or by storage of the compressed audio information for delivery by means of Short Messages.

In a preferred embodiment of the invention having this capability, illustrated in Fig. 3, ***the audio information is delivered 302 to the End User Audio Compression Function 304 of the End User***

Device 120, where it is encoded and compressed digitally. The End User Audio Compression Function 304 may be successfully implemented utilizing any of the methods or technologies discussed as being suitable for the implementation of the Message Center Audio Compression Function 110.

The compressed audio information is then formatted into one or more Short Messages by the End User Short Message Creation Function 306. The End User Short Message Creation Function 306 may be successfully implemented utilizing any of the methods or technologies discussed as being suitable for the implementation of the Message Center Short Message Creation Function 112. If the compressed audio information is too long to be contained in a single Short Message, it can be segmented into multiple Short Messages, utilizing any of the methods and technologies previously discussed. As before, the limit on the length of the total compressed audio information that may be processed at one time is an operational consideration that may vary depending on the facilities that will be used to deliver the Short Message(s) and on either user's subscription.

Once one or more Short Messages are created to contain the compressed audio information, the End User Short Message Delivery Function - Transmit 315 initiates transmission of the Short Message(s) to the Message Center 106 where they are received by the Message Center Short Message Delivery Function - Receive 325. Both the End User Transmit 315 and Receive 325

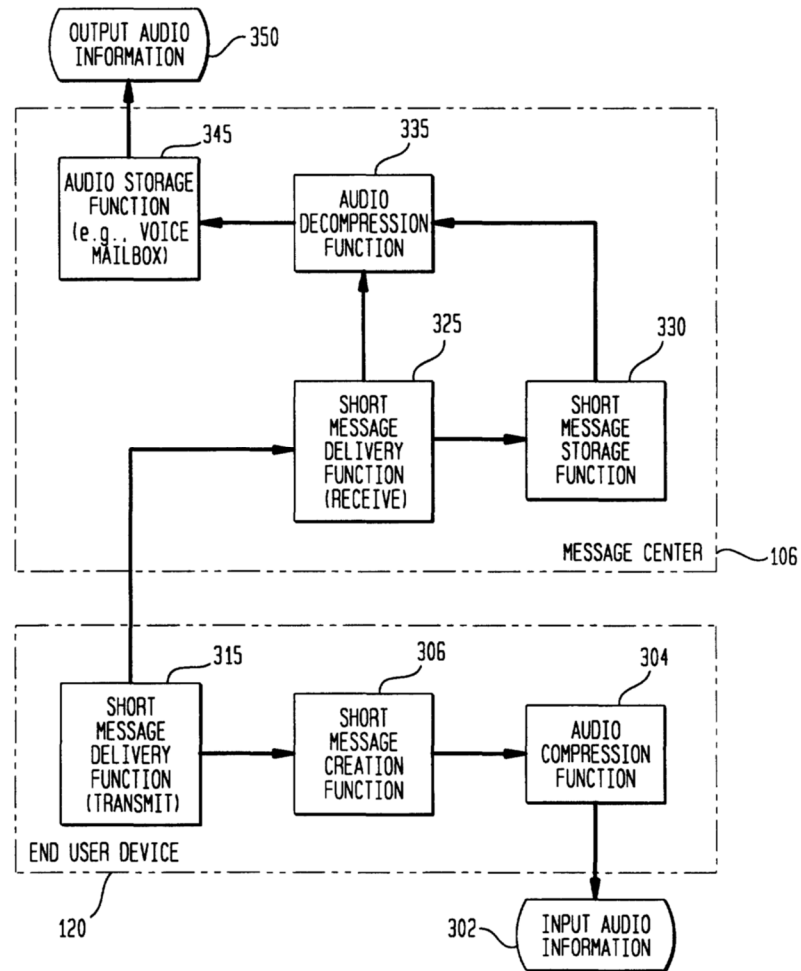
Functions may be successfully implemented utilizing any of the technologies and methods discussed for the Message Center Transmit 114 and Receive 121 Functions.

Upon successful reception of one or more Short Message(s), they are optionally stored locally by the Message Center Short Message Storage Function 330. *The compressed audio information may alternatively be passed immediately to the Message Center Audio Decompression Function 335. Both of these functions may similarly be implemented according to the technologies and methods previously discussed in conjunction with the End User Device.*

At a later time, the Short Message(s) may be either passed to the Message Center Audio Decompression Function 335, or may be *sent to another end user using the same techniques discussed in conjunction with the previous embodiments.* After the Message Center Audio Decompression Function 335, which reconstitutes the compressed audio information, the audio information is optionally stored by the Message Center Audio Storage Function 345 while awaiting delivery 350 to the intended recipient. Immediately, or at a later time, the reconstituted audio information may again be processed into compressed audio information using the techniques described in conjunction with the embodiments previously disclosed.

Id., 15:4-16:21.

FIG. 3

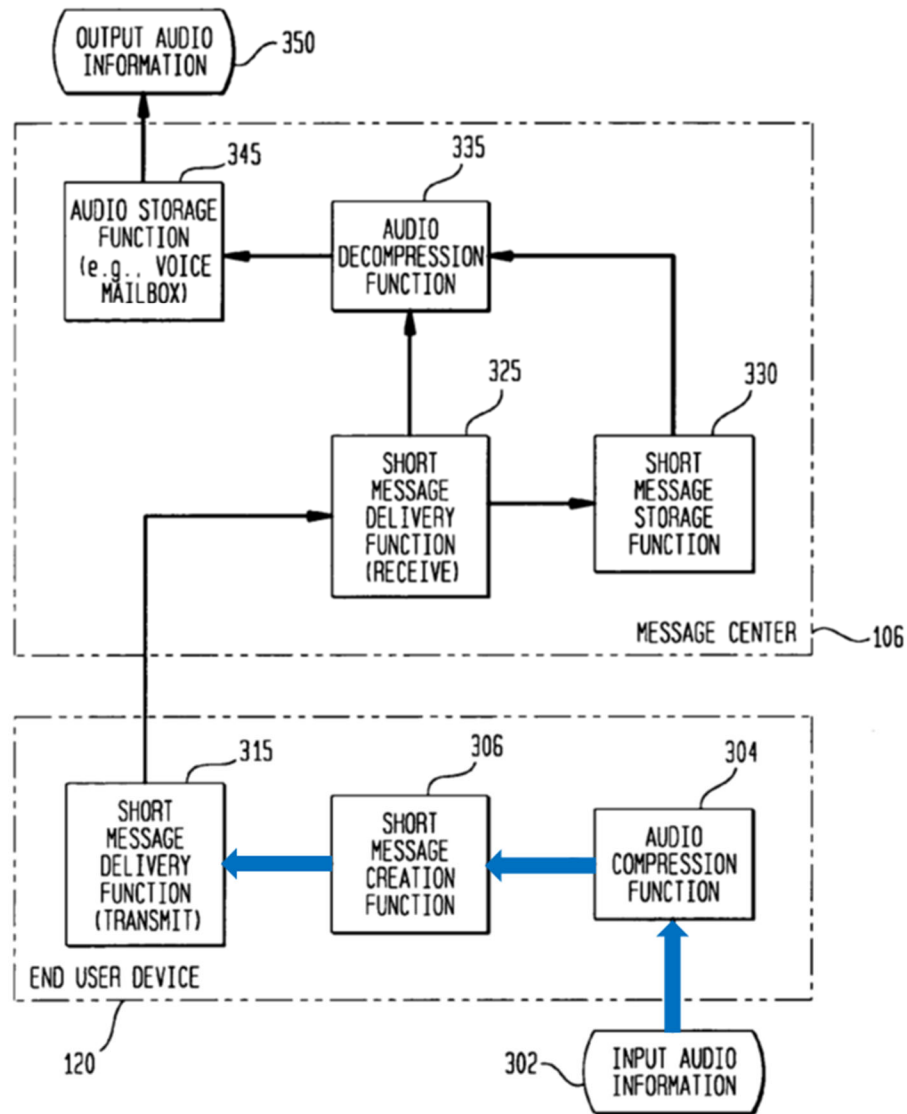


Ex. 1005, FIG. 3.

58. I note, however, that a person having ordinary skill in the art would have readily recognized that figure 3 includes a drafting error in that the direction of the flowlines with respect to the End User Device 120 should be in the opposite direction of what is shown. In my opinion, this error would not have impacted a person having ordinary skill in the art's understanding of

Dolan's disclosure. Figure 3, which is annotated below, shows the correct direction of the flow lines consistent with the description in the specification.

FIG. 3



Ex. 1005, FIG. 3 (annotated).

59. First, I note that the text on pages 15 and 16 of *Dolan* discusses figure 3 and explicitly and unambiguously explains that a user of an end user

device initiates sending (e.g., speaks) an audio message that is transmitted to another end user device. Ex. 1005, 15:4-7 (“The mobile user may also be the *initiator* of a short audio message utilizing this invention, either as a reply to a received message or independently.”). Moreover, consistent with the direction of flowlines in annotated figure 3 above, Dolan expressly describes that the audio information flows from Input Audio Information 302, to Audio Compression Function 304, to Short Message Creation Function 306, and to Short Message Delivery Function - Transmit 315:

- a. From Input Audio Information 302 to Audio Compression Function 304: “[I]llustrated in Fig. 3, the audio information is delivered 302 to the End User Audio Compression Function 304 of the End User Device 120, where it is encoded and compressed digitally.” *Id.*, 15:1-13.
- b. From Audio Compression Function 304 to Short Message Creation Function 306: “The compressed audio information is then formatted into one or more Short Messages by the End User Short Message Creation Function 306.” *Id.*, 15:17-18.
- c. From Short Message Creation Function 306 to Short Message Delivery Function - Transmit 315: “Once one or more Short Messages are created to contain the compressed audio information,

the End User Short Message Delivery Function - Transmit 315 initiates transmission of the Short Message(s) to the Message Center 106 where they are received by the Message Center Short Message Delivery Function - Receive 325.” *Id.*, 15:25-16-3.

60. Given *Dolan*’s written disclosure described above, a person having ordinary skill in the art would have understood that the information in the End User Device 120 shown in figure 3 necessarily flows from Input Audio Information 302, to Audio Compression Function 304, to Short Message Creation Function 306, to Short Message Delivery Function (Transmit) 315.

C. Freedman

61. *Freedman* is titled “Phonetic Speech-to-Text-to-Speech System and Method.” Ex. 1006, (54).



(19) United States		(10) Pub. No.: US 2004/0073423 A1	
(12) Patent Application Publication		(43) Pub. Date: Apr. 15, 2004	
Freedman			
(54) PHONETIC SPEECH-TO-TEXT-TO-SPEECH SYSTEM AND METHOD	(52) U.S. Cl.	704/235	
(76) Inventor: Gordon Freedman , Nepean (CA)	(57)	ABSTRACT	

Ex. 1006, Cover (annotated).

62. *Freedman* discloses a phonetic speech-to-text-to-speech system designed to optimize speech transmission with a small bandwidth from a source

to destination. *Id.*, Abstract (“A speech-to-text-to-speech for use with on-line and real time transmission of speech with a small bandwidth from a source to a destination.”); *see also id.* ¶[0008] (“Therefore, it is an object of the present invention to provide with ***a system that allows for on-line transmission of speech with a small bandwidth*** from a source to a destination.”)

63. *Freedman* further explains that speech is broken into phonemes and is transmitted via a telecommunication network:

Referring to FIG. 2, a schematic representation of a speech-to-text-to-speech system is shown. Speech is provided to a processor 20 through an input device in the form, for example, of a microphone 21. ***The processor 20 breaks the spoken words down into phonemes and translates the provided speech into a text that corresponds to the speech in a phonetic form. The phonetic text is transmitted via a telecommunication network such as the Internet or a public telephony switching system*** and is received by device 21 including a processor for restoring the speech according to the phonemes received. The restored speech is then provided to an output port 22 in the form for example of a loud-speaker.

Id. ¶[0035]; *see also id.* ¶[0048] (“Referring to FIG. 3b, the encoded transmitted text is received at input port 36 of the text-to-speech portion of the speech-to-text-to-speech system.”).

Referring to FIG. 5, a flow chart diagram of a method of using the speech-to-text-to-speech system is shown. An individual provides a speech to the system that breaks down the speech as it is provided to identify the phonemes in order to generate a text in a phonetic format. In a further step, the phonetic text is *sent through an existing communication system to a computer system remotely located. The transmission of the phonetic text is a fast process especially using Internet connections, and the texts such sent are usually small files transmitted.* Upon receiving the phonetic text, the speech generator on the remote computer system reconstitutes a speech based upon the phonetic system used.

Id. ¶[0063].

64. *Freedman* discloses that the system processes speech by breaking it into phonemes and “encoding the phonemes into series of symbols . . . other than a known symbolic representation of the speech in a known language,” thereby facilitating efficient transmission over telecommunication networks. *Id.* ¶[0011]; *see also id.* ¶¶[0020] (“encoding the phonemes into series of symbols compatible with a communication system, the series of symbols other than a known symbolic representation of the speech in a known language”), [0035] (“The phonetic text is transmitted via a telecommunication network.”), [0038] (“The encoding preferably results in a symbolic representation of the speech other than in a known human intelligible language.”), [0048] (“[T]he encoded

transmitted text is received at input port 36 of the text-to-speech portion of the speech-to-text-to-speech system.”), FIG. 5.

65. *Freedman* provides an example of a “resulting phonetic transform, **hə’ləʊðɪzənɪz’f aɪnɪŋ**,” which is encoded before being “transmitted through the output port 34 and received at the input port 36.” *Id.* ¶[0050]. As can be seen by this example, “the breakdown of a speech into symbols corresponding to known phonemes . . . renders the text unintelligible.” *Id.* ¶[0053].

D. Halonen

66. *Halonen* is titled “Methods and Apparatus for Updating the Software of a Mobile Terminal Using the Air Interface.” Ex. 1007, (54).



US005887254A

United States Patent [19]
Halonen

[11] **Patent Number:** **5,887,254**
[45] **Date of Patent:** **Mar. 23, 1999**

[54] **METHODS AND APPARATUS FOR
UPDATING THE SOFTWARE OF A MOBILE
TERMINAL USING THE AIR INTERFACE**

5,682,600 10/1997 Salin .
5,740,542 4/1998 Leeper et al. .
5,742,910 4/1998 Gallant et al. .
5,797,101 8/1998 Osmani et al. .

[75] Inventor: **Markku Halonen**, Oulu, Finland

FOREIGN PATENT DOCUMENTS

[73] Assignee: **Nokia Mobile Phones Limited**, Salo,
Finland

B-77395/91 5/1991 Australia .
0 275 510 A2 7/1988 European Pat. Off. .
0 459 344 A1 12/1991 European Pat. Off. .
0 478 231 A2 4/1992 European Pat. Off. .
2 204 973 11/1988 United Kingdom .
2 249 923 5/1992 United Kingdom .
2 256 734 12/1992 United Kingdom .

[21] Appl. No.: **638,126**

[22] Filed: **Apr. 26, 1996**

Ex. 1007, Cover (annotated).

67. *Halonen* is assigned to Nokia Mobile Phones Limited that, based on my personal knowledge, was one of the largest suppliers of mobile phones (also known as cellular phones) in the world during this timeframe. My personal phone in the 2004 timeframe was a Nokia phone that I used for voice calls as well as for sending and receiving SMS messages.

68. *Halonen* discloses the hardware components of a typical cellular phone, including a processor, a memory storing code, a microphone and a transceiver. For example, *Halonen* describes a “controller 18 . . . comprised of a digital signal processor (DSP) 18a and a microprocessor control unit (MCU) 18b,” to demodulate and decode received transmissions, including encoded voice, and perform high-speed signal processing. *Halonen*, 3:36-55.

69. *Halonen* explains the operation of the processor, stating that “[t]he MCU 18b . . . generally control[s] the operation of the terminal 10 to place and receive calls, to operate the keypad 22 and the display 20, to monitor the charge state of the battery 26, and other required functions.” *Id.*, 3:44-49.

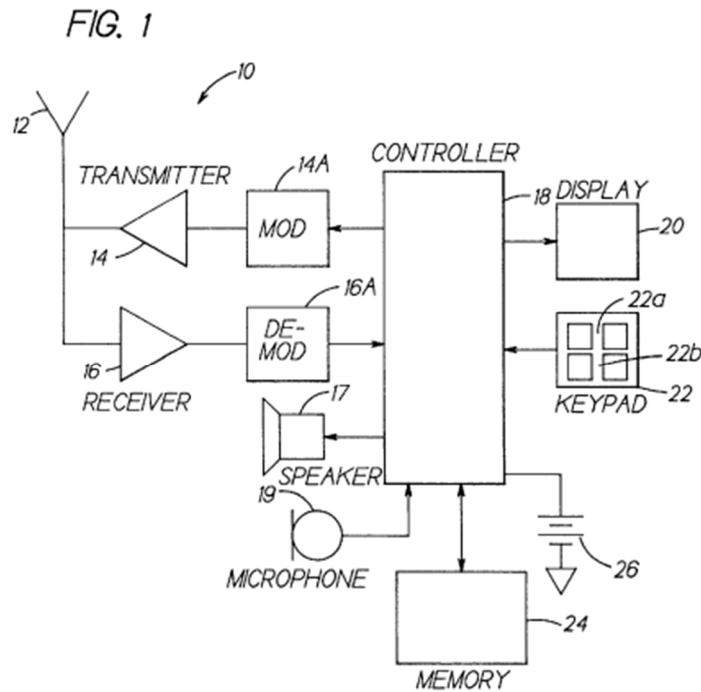
70. *Halonen* also discloses updating a mobile terminal’s operating program by receiving new or revised code blocks via Short Message Service (SMS) as follows, which are then acknowledged, processed, and stored for execution by the controller as follows:

As an example of the operation of the invention, assume that the MEM1 24b stores a currently valid operating program for the MCU 18b, and that the Active Memory flag is set to zero. The mobile terminal 10 then receives a message from the BMI 32 that a download of a new or revised operating program is to occur. Subsequently the BMI 32 begins transmitting the new or revised operating program to the mobile terminal 10 as a plurality of code blocks. ***By example, the code blocks can be sent using so-called Short Message Service (SMS) messages, or can be sent over a data channel, . . . The received code blocks are passed through the DSP 18a to the MCU 18b, which stores the received code blocks into the MEM2 24c.*** After the new or revised program is completely received, the MCU 18b changes the Active Memory flag in the NVM 24a to a one, thereby designating MEM2 24c as the Active memory and MEM1 24b as the Idle memory. ***The MCU***

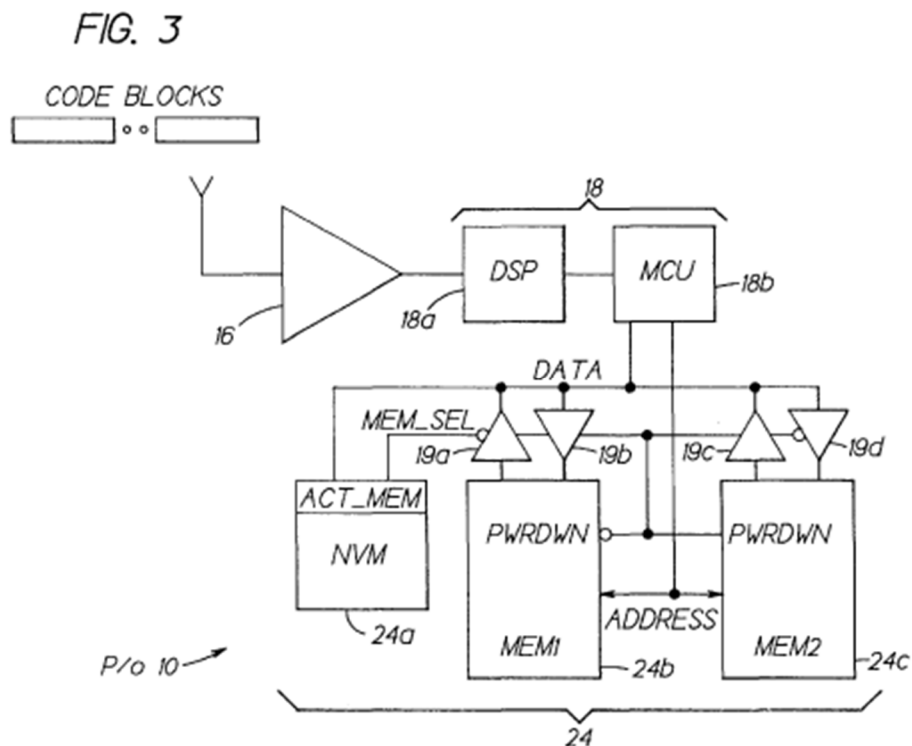
18b then resets itself to an initial state and begins executing code from the memory 24c.

Id., 4:47-5:12.

71. Further, as seen below in figures 1 and 3, *Halonen* depicts the processor (controller 18) comprising DSP 18a and MCU 18b that can receive, processes and executes code blocks.



Ex. 1007, FIG. 1.



72. The *Halongen* cellular phone has a “conventional microphone 19.” *Id.*, 2:61; see also *id.* FIG. 1.
73. The *Halongen* cellular phone also has a transceiver (for bidirectional RF communications with a network). *Id.*, Abstract (“A wireless user terminal (10), such as a cellular telephone, includes a transceiver (14, 16) for conducting bidirectional RF communications with a communications network (32).”), 1:61-65 (“A method of this invention is disclosed for operating a wireless mobile or user terminal having a transceiver for conducting bidirectional RF communications with a communications network.”), claim 1, FIG. 1.

74. The *Halonen* cellular phone also has a memory to store code for the controller. *Id.*, 3:7-9 (“An operating program for controlling the operation of controller 18 is also stored in the memory 24 (typically in a ROM device).”), Abstract (“The terminal further includes (a) a first memory capable of storing a data processor program; (b) a second memory capable of storing a data processor program.”), 2:11-13 (same); 3:64-66 (“The memories 24b and 24c are provided to store the operating program for the MCU 18b.”).

IX. CLAIM CONSTRUCTION

75. I have given all the claim terms their plain and ordinary meaning, as would have been understood by a person having ordinary skill in the art at the time of the alleged invention, which I understand is April 20, 2004, having taken into consideration the language of the claims, the specification, and the prosecution history of record.

X. SPECIFIC GROUNDS FOR CHALLENGE

76. In the sections below, I explain in detail how claims 1-17 of the '078 patent are not patentable over the prior art. More particularly:

77. **Ground 1A** challenges claims 1-3 and 8-11 as obvious over *Dolan*.

78. **Ground 1B** challenges claims 1 and 4-7 as obvious over the combination of *Dolan* and *Freedman*.

79. **Ground 2A** challenges claims 12 and 15-17 as obvious over the combination of *Dolan* and *Halonen*.

80. **Ground 2B** challenges claims 12-14 as obvious over the combination of *Dolan*, *Halonen*, and *Freedman*.⁴

A. Ground 1A: *Dolan* Renders Obvious Claims 1-3 and 8-11

1. Claim 1

a. 1[pre]: “A method of sending a voice message via a mobile communication device, said method comprising:”

81. I understand that “[a] method of sending a voice message via a mobile communication device, said method comprising” is the preamble to claim 1 of the ’078 patent. I have been asked to assume that the preamble is a claim limitation. Under that assumption, in my opinion, *Dolan* discloses or suggests “a method of sending a voice message via a mobile communication device.”

⁴ To the extent my declaration refers to prior art references other than those recited in the above-listed grounds 1A, 1B, 2A, and 2B, I am using such references as evidence supporting what I believe would have been known to a person of ordinary skill in the art at the time of the alleged invention, for example, about the state of the art and/or the benefits and advantages of certain design choices.

82. For example, *Dolan* discloses:

The present invention relates to *delivery of audio messages* to or from remote receiving units such as *mobile telephones* and, in particular, to a method for using Short Message Service or other discrete message-oriented services to deliver audio messages.

Ex. 1005, 1:4-6.

83. *Dolan* further discloses the need for an end user device (or “remote receiving unit”) to receive and send audio messages via SMS, stating, “[w]hat has been needed, therefore, is a way to conveniently send audio messages directly to and from the user of a remote receiving unit.” *Id.*, 2:26-27. Moreover, a “mobile user may also utilize this invention to be the initiator of a short audio message to a second end user device, with the functions being implemented in a manner analogous to that for Message Center to End User transmission.” *Id.*, 5:21-23.

84. *Dolan* further discloses:

The compressed audio information is next transmitted by means of Short Message Service or another discrete message-oriented service to the end user. The Short Message containing the compressed audio information is received at an end user device, the audio information is decompressed, and the decompressed audio information is played back to the end user (the intended message recipient). The transmission scheme used by the

preferred embodiment of this invention is Short Message Service

...

Id., 7:28-8:3.

At the time that the End User Device 120 becomes available for delivery, a Message Center Short Message Delivery Function - Transmit 114 initiates transmission of the Short Message(s) to the End User Device 120, where they are received by an End User Short Message Delivery Function - Receive 121. This function is standard to SMS, but the specific implementation used is actually determined by the type of receiving device, the transmission medium, and the service protocol being used. For example, if the message is being delivered to a mobile unit, the SMS protocol might be used . . .

Id., 12:18-24.

The mobile user may also be the initiator of a short audio message utilizing this invention, either as a reply to a received message or independently. As illustrated in Fig. 3, the flow of the audio information is from the End User Device 120 to the Message Center 106 for subsequent delivery to a second end user.

Id., 15:4-7.

85. *Dolan* contemplates a preferred embodiment in which an “end user device is conceived as a digital mobile telephone handset.” *Id.*, 4:12-13; *see also id.*, 10:22-23 (“The end user device is conceived as a digital mobile telephone handset.”).

86. Further, a person having ordinary skill in the art would have understood that the audio messages discussed in *Dolan* encompass *voice* messages. For example, *Dolan* provides the explicit example of an audio message that includes the phrase “Call home, emergency!” (*Dolan*, 6:16-17), which a person having ordinary skill in the art would have understood would have been spoken by a person and therefore is a human voice.

87. *Dolan* also specifically references a “spoken phrase.” *Id.*, 10:27-28 (“[I]n a preferred embodiment of the system the audio information 102, *such as a spoken phrase*, is delivered to the Message Center Audio Compression Function 110 within a Message Center 106, where it is encoded.”)

88. Additionally, *Dolan* discloses encoding audio messages using “technology used in *voice* coding schemes on wireless networks” (*id.*, 7:9-12), which a person having ordinary skill in the art would also have understood to mean that *Dolan*’s audio messages encompass voice messages.

89. Accordingly, a person having ordinary skill in the art would have understood *Dolan* discloses or suggests sending an audio message (“voice message”) via an end user device such as a digital mobile telephone handset (“via a mobile communication device”).

b. 1[a]: “receiving an utterance from a user of the mobile communication device;”

90. In my opinion, *Dolan* discloses or suggests “receiving an utterance from a user of the mobile communication device.”

91. For example, *Dolan* has an embodiment (describing and shown in figure 3⁵), in which the user of an end user device initiates sending a short audio message. Ex. 1005, 15:4-5 (“The mobile user may also be the initiator of a short audio message utilizing this invention.”).

92. As I discussed in Section X.A.1.a, *Dolan* discloses that the end user device may be a digital mobile telephone handset. *Dolan* also discloses that the digital mobile telephone handset receives the audio message:

As illustrated in Fig. 3, the flow of the audio information is from the End User Device 120 to the Message Center 106 for subsequent delivery to a second end user. . . . In a preferred embodiment of the invention having this capacity, illustrated in Fig. 3, the audio information is delivered 302 to the End User Audio Compression Function 304 of the End User Device 120, where is encoded and compressed digitally.

Id., 15:5-13.

⁵ While *Dolan*’s figure 3 includes a drafting error, as I discussed above, this would not have impacted a person having ordinary skill in the art’s understanding of *Dolan*’s disclosure.

93. *Dolan* also discloses that the audio message can be a user speaking words, for example, “Call home, emergency!” *Id.*, 6:16-17. The plain and ordinary meaning of “utterance” is “a spoken word, statement, or vocal sound.” Ex. 1008. As such, in my opinion, a person having ordinary skill in the art would have understood “Call home, emergency!” to constitute an utterance.

94. Therefore, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests a digital mobile telephone handset receiving an audio message (“receiving an utterance”) spoken by a user of the device (“from a user of the mobile communication device”).

c. 1[b]: “generating a non-text representation of the received utterance;”

95. In my opinion, *Dolan* discloses or suggests “generating a non-text representation of the received utterance.”

96. For example, *Dolan* discloses that “audio information is delivered 302 to the End User Audio Compression Function 304 of the End User Device 120, ***where it is encoded and compressed*** digitally.” Ex. 1005, 15:12-13. Additionally, *Dolan* discloses that there are various possible compression techniques, and “[a]s a general proposition, in each of these compression techniques a voice signal is prepared for digital transmission by band-limiting and sampling at a convenient pre-specified frequency (such as 8 kHz).” *Id.*, 7:17-19.

97. *Dolan* discloses that the Message Center Audio Compression Function 110 may be implemented “by any of the devices and methods known in the art that have the capability of encoding and compressing audio signals to produce a digital signal as an output.” *Id.*, 11:6-8.

98. *Dolan* goes on to say that the end user device can compress audio using the same methods or technologies as the message center. Specifically, *Dolan* states that the “End User Audio Compression Function 304 may be successfully implemented utilizing any of the methods or technologies discussed as being suitable for the implementation of the Message Center Audio Compression Function 110.” *Id.*, 15:14-16.

99. One such method for converting normal speech to a “compressed digital” form (as *Dolan* explains) is to utilize various vocoders already utilized by mobile phones in the industry:

The input begins as a normal speech path and is converted to compressed digital via devices internal to the Message Center that perform the compression function 110. Examples are any of the various vocoders utilized by mobile phones in the industry already, such as the GSM Full-Rate Voice Coder, GSM Half-Rate Voice Coder, GSM Enhanced Full-Rate Voice Coder, 13KB IS95 Voice Coder, 8 KB IS95 Voice Coder, EVRC Voice Coder, and the IS 13 6 8KB Voice Coder.

Id., 11:8-14.

100. As I have noted in the preceding paragraphs, *Dolan* discloses the compression of audio information using voice coding schemes/technology such as known vocoders that were already used by mobile phones in the industry. In my opinion, a person having ordinary skill in the art would have also understood that *Dolan*'s compressed audio information is a non-text representation of the voice/speech because it is a compressed version of an audio signal that is not human readable text.

101. *Dolan*'s disclosure of compressed audio information is similar to how the '078 patent's disclosure which recognizes that audio information compressed using a vocoder is a non-text representation of the audio. Ex. 1001, 6:58-61 ("[I]nstead of using a phoneme recognizer, a vocoder could be used to generate the compressed representation and then that compressed representation would be inserted into the body of the SMS message." Ex. 1001, 6:58-61.

102. Therefore, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests encoding and compressing ("generating a non-text representation") an audio signal such as "Call home, emergency!" received by the digital mobile telephone handset ("of the received utterance").

d. 1[c]: “inserting at least some of the non-text representation into a body of a text message; and”

103. In my opinion, *Dolan* discloses or suggests “inserting at least some of the non-text representation into a body of a text message.”

104. For example, *Dolan* discloses that the compressed audio information is then formatted “into one or more Short Messages by the End User Short Message Creation Function 306.” Ex. 1005, 15:17-18; *see also id.*, 3:14-15 (“The compressed audio information is transmitted by means of Short Message Service or another discrete message-oriented service to the end user.”), 7:26-27 (“The compressed audio information is next transmitted by means of Short Message Service or another discrete message-oriented service to the end user.”), 8:2-3 (“The transmission scheme used by the preferred embodiment of this invention is Short Message Service.”).

105. As *Dolan* explains, the standard implementation of Short Message Service (before implementing the enhancements described in *Dolan*) is used to send and receive text messages on digital mobile phones. *See, e.g., id.*, 1:18-19 (“SMS is currently utilized only to deliver messages to mobile stations as short *text* messages.”) (emphasis in original), 2:20-21 (“Short Message Service capability is two-way, i.e. a digital mobile phone user can both receive and send short *text messages* utilizing this technology.”), 3:1-3 (“In particular, an object

of the present invention is to utilize Short Message Service or other existing *text message-oriented services* to deliver short audio messages.”).

106. *Dolan*’s disclosure is consistent with the ’078 patent, which identifies SMS messages as a form of text messages. Ex. 1001, 2:31-32 (stating “the text message is an SMS message”).

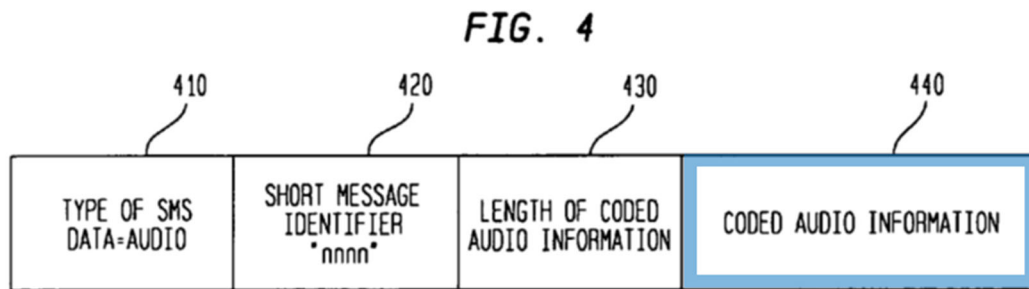
107. *Dolan* also discloses that “[t]he End User Short Message Creation Function 306 may be successfully implemented utilizing any of the methods or technologies discussed as being suitable for the implementation of the Message Center Short Message Creation Function 112.” Ex. 1005, 15:18-21.

108. *Dolan* discloses that “the Message Center Short Message Creation Function” takes the “compressed audio information” and formats it “into one or more Short Messages.” *Id.*, 4:20-21. Similarly, *Dolan* also states:

The compressed audio information is then formatted into one or more Short Messages by the Message Center Short Message Creation Function 112. The Message Center Short Message Creation Function 112 is implemented by any of the devices and methodologies know [sic] in the art for implementing a Short Message or other discrete message-oriented service, such as SMS, where a header and length field wrap the actual message during transmission. Examples of these are found in the TDMA, CDMA and GSM standards.”

Id., 11:15-21.

109. Additionally, with reference to figure 4, *Dolan* describes the major fields in a standard SMS message, including the coded audio information 440 to contain the encoded, compressed audio information. Specifically, *Dolan* discloses that the following SMS major fields “used to carry coded audio information are a Type of Data field 410 . . . , a Short Message Identifier field (‘nnnn’) 420 . . . , a Length field 430 . . . , and the coded audio information 440.” *Id.*, 8:14-20.



Ex. 1005, FIG. 4 (annotated).

110. *Dolan*’s figure 4 embodiment “assumes that all of the coded audio information 440, plus the Type of Data 410, Identifier 420, and Length 430 fields, will fit into a single SMS message sent across one or more intervening networks between a Message Center and the end user’s equipment (e.g., a mobile phone, PC, etc.).” *Id.*, 8:20-24.

111. A person having ordinary skill in the art would have understood coded audio information field 440 to comprise a body of the SMS message

because it contains the encoded, compressed audio information that is being transmitted by the SMS message.

112. At a minimum, a person having ordinary skill in the art would have understood the encoded, compressed audio information to have been inserted into the body of the SMS message because fields 410-440 together comprise the SMS message where the audio message is block 440, and that audio message is at least part of the SMS body.

113. Accordingly, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests inserting the compressed audio information (“inserting at least some of the non-text representation”) into coded audio information field 440 of a SMS message (“into a body of a text message”).

e. 1[d]: “sending the text message including the at least some of the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device.”

114. In my opinion, *Dolan* discloses or suggests “sending the text message including the at least some of the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device.”

115. *Dolan* explains that “[t]he *mobile user* may also be the initiator of a short audio message utilizing this invention.” Ex. 1005, 15:4-5. As *Dolan*

explains with reference to figure 3, in such circumstances “[t]he flow of the audio information is from the End User Device 120 to the Message Center 106 for subsequent delivery to a second end user.” *Id.*, 15:5-7; *see also id.*, 5:21-22 (“The mobile user may also utilize this invention to be the initiator of a short audio message to a second end user device.”).

116. *Dolan* describes how the audio message is sent. Specifically, *Dolan* discloses that the:

End User Short Message Delivery Function – Transmit 315 initiates transmission of the Short Message(s) [containing the compressed audio information] to the Message Center 106 where they are received by the Message Center Short Message Delivery Function – Receive 325.

Ex. 1005, 16:1-3; *see also id.*, 16:3-6 (“Both the End User Transmit 315 and Receive 325 Functions may be successfully implemented utilizing any of the technologies and methods discussed for the Message Center Transmit 114 and Receive 121 Functions.”).

117. *Dolan* states that “the Short Message(s) may be either passed to the Message Center Audio Decompression Function 335, *or* may be sent to another end user using the same techniques discussed in conjunction with the previous embodiments.” *Id.*, 16:13-15. Because *Dolan* discloses these options as alternatives, a person having ordinary skill in the art would have understood

Dolan to disclose that the message need not be decompressed before it is sent to another end user.

118. *Dolan* discloses that “[t]he transmission scheme used by the preferred embodiment of this invention is Short Message Service” *Id.*, 8:2-3. In my opinion, a person having ordinary skill in the art would have understood that communication of SMS messages (e.g., from an end user device and to the *recipient’s* end user device that receives the audio message) occur over a wireless messaging channel.

119. For example, *Dolan* discloses that Short Message Service is “currently defined in various standards such as GSM (Global System for Mobile Communications).” *Id.*, 1:21-22; *see also id.*, 11:16-21 (“The Message Center Short Message Creation Function 112 is implemented by any of the device and methodologies know [sic] in the art for implementing a Short Message or other discrete message-oriented service, such as SMS, where a header and length field wrap the actually message during transmission. Examples of these are found in the TMDA, CDMA and GSM standards.”).

120. A person having ordinary skill in the art would have known that GMS, TDMA and CDMA were different standards for wireless communication implemented by mobile telephone devices that were commonly known and referred to as “cellular phones.” For example, U.S. Patent No. 5,870,427, titled

“Method for multi-mode handoff using preliminary time alignment of a mobile station operating in analog mode,” (hereinafter “*Tiedemann*”) discloses the field of cellular communications and identifies that the GSM, TDMA, and CDMA are standards within the field:

In the field of *cellular communications*, systems employing both analog and digital systems have been developed. . . . Systems operating in the digital mode may be of different types of digital modulation such as *time division multiple access system (TDMA)* or a spread spectrum multiple access type, employing a frequency hopping (FH) or direct sequence spread spectrum modulation mode. Digital multiple access systems employing direct sequence spreading are commonly referred to as *Code Division Multiple Access (CDMA)* systems. An example of a TDMA system is one that is described in the Telecommunications Industry Association/Electronic Industry Association Interim Standard TIA/EIA/IS-54 “*Cellular System* Dual-Mode Mobile Station-Base Station Compatibility Standard”, and the revisions thereto (TIA/EIA/IS-54A and TIA/EIA/IS-54B). Another example of a TDMA system is the system known as the *Global System for Mobile Communication (GSM)*. An example of a CDMA system is one that is set forth in the Telecommunications Industry Association/Electronic Industry Association Interim Standard TIA/EIA/IS-95 “Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum *Cellular System*”, and the revisions thereto (TIA/EIA/IS-95A). Although

these standards are typically defined for frequency spectrum specific to *cellular communications*, the analog and digital techniques are also applicable to other frequencies such as those defined for personal communication services (PCS).

Ex. 1009, 1:20-52.

121. Similarly, U.S. Patent No. 6,081,724, titled “Portable communication device and accessory system,” (hereinafter “*Wilson*”), discloses GSM, TDMA, and CDMA as they relate to wireless cellular systems:

There are presently multiple types of *cellular radiotelephone systems* operating. These systems include the frequency modulated (FM) advanced mobile phone system (AMPS) and *two digital cellular systems: time division multiple access (TDMA and GSM), and code division multiple access (CDMA)*. The digital cellular systems are being implemented to handle capacity problems that AMPS is experiencing. Dual-mode CDMA/FM radiotelephones exist which are selectively operative in either FM or CDMA modes. Telecommunications Industry Association (TIA)/Electronic Industries Association (EIA) Interim Standard 95, entitled “Mobile Station-Base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System” sets forth the requirements and standards for a dual-mode radiotelephone which is selectively operative in either FM or CDMA modes.

Ex. 1010, 1:13-29.

122. As another example, U.S. Patent No. 5,963,548, titled “Apparatus and method for configuring a data channel for symmetric/asymmetric data transmission,” (hereinafter “*Virtanen*”), discusses GSM, TDMA, and CDMA and their role in the cellular industry:

Several types of technologies have become dominant in the *cellular industry*. In the United States most *cellular systems* currently operating use analog signal transmission techniques, as specified by the Telecommunications Industry Association/Electronic Industry Association(TIA/EIA) AMPS standard, or a combination of analog and time division multiple access (TDMA) signal transmission techniques, as specified by the TIA/EIA IS-54 and IS-136 standards. In Europe, *cellular systems* may operate according to one of several analog system standards, depending on the country, or according to the digital Global Services for Mobile (GSM) TDMA standard that has been specified for Europe. In other parts of the world most cellular systems operate according to one of the standards used in the United States or Europe, except for in Japan where the TDMA personal digital communication (PDC) standard has been developed and is in use. However, despite the present dominance of analog and TDMA technologies, the *cellular industry* is dynamic and new technologies are constantly being developed as alternatives to these currently dominant technologies. One alternative digital signal transmission technique that has recently

been the focus of attention for cellular systems is known as code division multiple access (CDMA).

Ex. 1011, 1:32-55.⁶

123. In my opinion, a person having ordinary skill in the art would have known that GSM, TDMA and CDMA were different standards for wireless communication implemented by mobile telephone devices that were commonly known and referred to as “cellular phones.”

124. Additionally, *Dolan* explicitly states that it uses wireless technology. For example, *Dolan* states that “[t]he Message Center 106 will commonly be some form of voice mailbox system that provides messaging services and/or paging to the user and is attached to the intervening network by any of the variety of export techniques known in the art for providing end-to-end delivery of digital data, including, but not limited to: Internet, signaling systems, and *wireless* (radio, etc.).” Ex. 1005, 11:1-5; *see also id.*, 8:2-7 (“The transmission scheme used by the preferred embodiment of this invention is Short Message Service or any other discrete message-oriented service wherein Short

⁶ To be clear, here I am relying on Ex. 1009 (*Tiedemann*), Ex. 1010 (*Wilson*), and EX. 1011 (*Virtanen*) as evidence of a person having ordinary skill in the art’s knowledge and the state of the art. These references are not part of the unpatentability grounds.

Messages are captured on behalf of an end user, stored in a messaging center until the end user device is available to receive delivery, and then delivered digitally to the end user, such as the user of a digital *wireless telephone device*.”)

125. Also, *Dolan*’s disclosure in this regard is the same or similar to the ’078 patent’s disclosure, which states: “***Most legacy wireless network systems such as GSM, TDMA, and CDMA have a text/data channel capable of sending and receiving SMS***, so the infrastructure for this service already exists even in emerging markets in developing countries.” Ex. 1001, 1:50-54. Further, the ’078 patent states “[t]he wireless messaging channel is an SMS channel . . .” (*id.*, 2:30-31), and the “SMS application 180 then wirelessly transmits the SMS message over the SMS channel to the recipient’s cell phone,” (*id.*, 4:62-64).

126. While it is my opinion that *Dolan* discloses this feature for the reasons I have stated above, Petitioners’ counsel has asked me to consider if the Patent Owner argues that *Dolan*’s disclosure does not disclose this feature because *Dolan*’s figure 3 does not explicitly show the Message Center sending the Short Message to another End User Device. I have evaluated this argument, and to the extent Patent Owner raises it, it is not correct.

127. Contrary to this first argument, in my opinion, a person having ordinary skill in the art would have understood *Dolan* as a whole to disclose or suggest that a user of an end user Device can send an audio message to another

end user device using one or more Short Messages (e.g., SMS messages). For example, *Dolan* explicitly discusses that “[t]he mobile user may also utilize this invention to be the initiator of a short audio message **to a second end user device.**” *Dolan*, 5:21-22. Regarding figure 3 specifically, *Dolan* states that “the flow of the audio information is from the End User Device 120 to the Message Center 106 for subsequent delivery **to a second end user.**” *Id.*, 15:5-7.

128. In my opinion, the fact that figure 3 does not explicitly show the second end user device would not have altered a person having ordinary skill in the art’s understanding of *Dolan* as whole, especially in view of its written description.

129. Petitioners’ counsel has asked me to consider if the Patent Owner argues that *Dolan* requires that the compressed audio information in the Short Message is decompressed before it is sent by the Message Center to another End User Device. Contrary to this argument, in my opinion, a person having ordinary skill in the art would have understood *Dolan*’s disclosure to encompass an embodiment in which, in the context of one end user device sending an audio message to another end user device using SMS, the Message Center receives the SMS message from the first end user device but does not decompress the audio information before forwarding it to the second end user device.

130. For example, *Dolan* discloses that, in the figure 3 embodiment, the audio information is delivered to the Message Center “for subsequent delivery to a second user” and that this “subsequent delivery may be accomplished [1] by reconstitution of the compressed audio information [i.e., decompression] and optional storage in a voice mailbox, or [2] by storage of the compressed audio information for delivery by means of Short Messages.” Ex. 1005, 15:7-10. In my opinion, a person having ordinary skill in the art would have understood that the second option involves delivery of the “compressed audio information” by a text message.

131. Thus, in my opinion, a person having ordinary skill in the art would have understood *Dolan* to encompass an embodiment in which, when a first end user device sends an audio message using SMS to a second end user device, *Dolan*’s Message Center does not decompress the SMS message, but instead forwards it to the second end user device that performs the decompression.

132. Even if *Dolan* did not expressly or inherently encompass the forwarding a short message to another end user device without first decompressing it (which for the reasons already discussed, it does), *Dolan* still discloses or suggests this limitation. Nothing in the claim language, precludes decompression/recompression of audio information before the final end user received the SMS message. Claim limitation 1[d] is only directed to what is *sent*

by the first end user device (“mobile communication device”), not what is received by the recipient’s device.

133. Additionally, in my opinion, even if *Dolan*’s written description suggested the Message Center decompresses the audio information contained in every Short Message (and it does not), it would have been obvious to a person having ordinary skill in the art to modify the functionality of *Dolan*’s Message Center such that audio information contained in Short Messages that are received but are intended for delivery to another end user device are not decompressed.

134. In my opinion, a person having ordinary skill in the art would have readily appreciated the advantages of such a modification, including that (i) as the Message Center has no need to replay the audio information, there is no reason for the audio information to be decompressed and (ii) not decompressing the audio information would avoid unnecessarily expending processing resources to decompress the audio information.

135. Therefore, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests sending compressed audio information that has been inserted into an SMS message (“sending the text message including the at least some of the non-text representation”) using an SMS protocol incorporated into at least one of the GSM, TDMA or CDMA

wireless communication standards (“over a wireless messaging channel”) from one digital mobile telephone handset (“from the mobile communication device”) to a second digital mobile telephone handset (“to a recipient’s device”).

2. Claim 2: “The method of claim 1, wherein the mobile communication device is a cellular phone.”

136. In my opinion, *Dolan* discloses or suggests “wherein the mobile communication device is a cellular phone.”

137. For example, *Dolan* discloses that “in the preferred embodiment the end user device is conceived as a digital mobile telephone handset.” Ex. 1005, 4:12-13, 10:22-23. Additionally, *Dolan* refers to the end user device as a “mobile telephones,” (*id.*, 1:5, 3:5-6, 6:14), “mobile phone,” (*id.*, 6:20), or “wireless telephone device,” *id.*, 1:27-2:1.

138. *Dolan* explains that its “invention is easily codified into mobile phone standards” (*id.*, 7:3), and that it uses the SMS protocol to send audio messages to digital mobile telephone handsets, stating:

In Short Message Service, as currently defined in various standards such as GSM (Global System for Mobile Communications) [see, e.g., TIA/EIA/IS-637, Short Message Services for Wide Band Spread Spectrum ***Cellular Systems***, 1995; TIA/EIA/TSB-79, Short Message Services for Wide Band Spread Spectrum ***Cellular Systems***, 1997 (an update to IS-637); ***TS GSM 03.40***, Digital ***cellular telecommunications system***, Technical

realization of the Short Message Service (SMS); Point-to-Point (PP)] . . .

Id., 1:21-26.

139. *Dolan* further states that examples “for implementing a Short Message or other discrete message-oriented service, such as SMS” could be “found in the TDMA, CDMA and GSM standards.” *Id.*, 11:18-21.

140. In my opinion, a person having ordinary skill in the art would have known that digital mobile telephone handsets or mobile telephones that implement the TDMA, CDMA and/or GSM standards were commonly known and referred to as cellular telephones. *See* also my discussion of these standards associated with cellular phones in Section X.A.1.e. (also discussing Ex. 1009 (*Tiedemann*), Ex. 1010 (*Wilson*), and Ex. 1011 (*Virtanen*)).⁷

141. *Dolan*’s disclosure is similar to the ’078 patent, which identifies mobile phones that implement the GSM standard as cellular phones:

In the described embodiment, the ***cellular phone*** is a ***smartphone 300***, such as is illustrated by the high-level functional block diagram of FIG. 3. Smartphone 300 is a Microsoft PocketPC-powered phone which includes at its core a baseband DSP 302 (digital signal processor) for handling the cellular communication functions (including for example voiceband and channel coding functions) and an applications processor 304 (e.g. Intel StrongArm

⁷ *See supra* note 6.

SA-1110) on which the PocketPC operating system runs. The phone ***supports GSM voice calls***, SMS (Short Messaging Service) text messaging, wireless email, and desktop-like web browsing along with more traditional PDA features.

Ex. 1001, 7:38-46.

142. Accordingly, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests that the digital mobile telephone handset (“mobile communication device”) can implement any one of known standards such TDMA, CDMA and/or GSM (“is a cellular phone”).

3. Claim 3: “The method of claim 1, wherein inserting at least some of the non-text representation involves inserting all of the non-text representation into the body of the text message.”

143. In my opinion, *Dolan* discloses or suggests “wherein inserting at least some of the non-text representation involves inserting all of the non-text representation into the body of the text message.”

144. As I explained above in Section X.A.1.d, *Dolan* discloses an embodiment in which all the compressed audio information 440, plus the Type of Data 410, Identifier 420, and Length fields, will fit into a ***single SMS message***.

145. Therefore, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests an embodiment in

which inserting the compressed audio information into the SMS message (“inserting at least some of the non-text representation”) involves inserting all of the compressed audio information (“involves inserting all of the non-text representation”) into the body of the SMS message (“into the body of the text message”).

4. Claim 8: “The method of claim 1, wherein the wireless messaging channel is an SMS channel and the text message is an SMS message.”

146. *Dolan* discloses or suggests this claim for the reasons I discussed in Section X.A.1.e (“the wireless messaging channel is an SMS channel”), and Section X.A.1.d (“the text message is an SMS message”).

5. Claim 9: “The method of claim 1, further comprising including an indicator with the text message identifying the text message as containing a non-text representation of the utterance.”

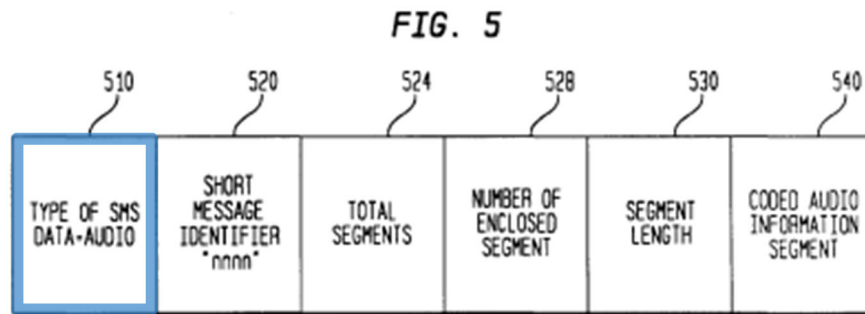
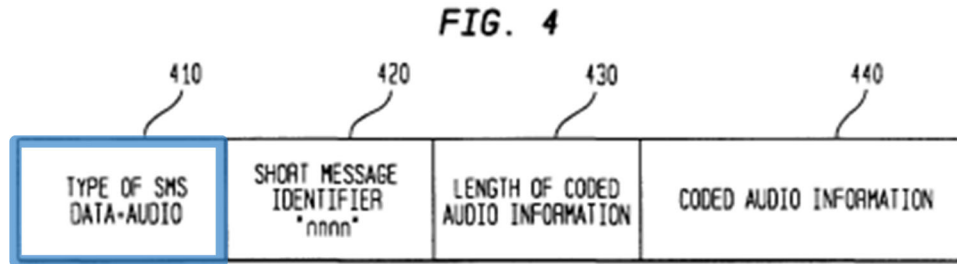
147. In my opinion, *Dolan* discloses or suggests “including an indicator with the text message identifying the text message as containing a non-text representation of the utterance.”

148. For example, *Dolan* discloses that “[t]he Short Message containing all or portions of the encoded audio information also contains a *type field* that indicates the contents of the Short Messages, as well as the type of audio decompression technique necessary to reconstitute the audio information.” Ex.

1005, 8:11-13. *Dolan* shows this type filed as “Type of Data field 410” (*id.*, 8:15, figure 4), or “Type of Data field 510” (*id.*, 10:7-8, figure 5).

149. *Moreover*, in the description for figure 4, *Dolan* states that the “Type of Data field 410, indicat[es] that the associated message is an Audio message and include[es] information on the type of translator needed to decompress and decode the message field in order to retrieve the audio message.” *Id.*, 8:14-17. In the description for figure 5, *Dolan* explains that the “Type of Data field 510 is the same as for a non-segmented message.” *Id.*, 10:7-8.

150. As can be seen in figures 4 and 5 (depicted below), each shows the content of the Type of Data field being as “Type of SMS Data = Audio.” As I explained above in §X.A.1.c, the compressed audio information to which “Type of SMS Data = Audio” refers is a non-text representation of the audio message.



Ex. 1005, FIGs. 4 & 5 (annotated).

151. Therefore, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests including a Type of Data field with the SMS message (“including an indicator with the text message”) which indicates the associated message is a compressed audio of the user speaking the words, for example, “Call home, emergency!” (“identifying the text message as containing a non-text representation of the utterance”).

6. Claim 10: “The method of claim 1, wherein the non-text representation is a compressed version of the received utterance.”

152. In my opinion, *Dolan* discloses or suggests this claim for the same reasons I discussed above in Section X.A.1.c.

7. Claim 11

- a. *11[a]: “The method of claim 1, further comprising dividing the non-text representation into sections and inserting each section into a body of a different corresponding text message of a plurality of text messages, wherein the first-mentioned text message is one of said plurality of text messages and”*

153. In my opinion, *Dolan* discloses or suggests “dividing the non-text representation into sections and inserting each section into a body of a different corresponding text message of a plurality of text messages, wherein the first-mentioned text message is one of said plurality of text messages.”

154. With reference to figure 3, *Dolan* explains audio information that is too long can be segmented and placed into multiple Short Messages:

If the compressed audio information is too long to be contained in a single Short Message, it can be segmented into multiple Short Messages, utilizing any of the methods and technologies previously discussed.

Ex. 1005, 15:21-24.

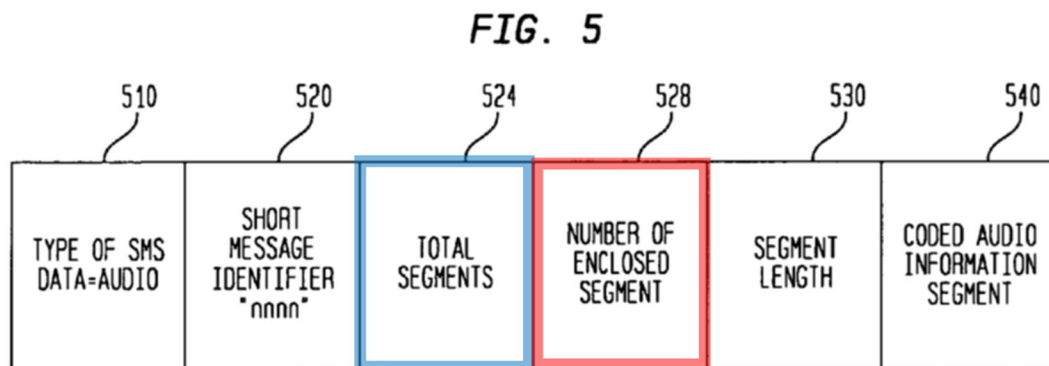
155. Additionally, *Dolan* explains that in segmenting messages, “[t]he segmentation and reassembly process [can be] carried out by the two endpoints which are transmitting and receiving the coded audio information” *Id.*, 9:26-28.

156. More particularly, *Dolan* discloses that “[c]oded audio information that is too long for a single Short Message is broken into segments, each of

which is placed into a separate Short Message having the fields: Type of Data; Short Message Identifier; Total Segments; Number of Enclosed Segment; Segment Length; and Coded Audio Information Segment.” *Id.*, 4:1-5. *Dolan* also states (as shown in figure 5), that “coded audio information that is too long for a single Short Message is broken into segments, for example, into Segment 1 of 3, Segment 2 of 3, and Segment 3 of 3. Each of these segments is then placed into a separate Short Message” *Id.*, 10:1-3.

157. *Dolan* also discloses that when audio information is broken into multiple segments, “[e]ach Short Message h[as] an indicator signaling that the segment contained within the Short Message is a particular segment number out of some total number of segments” *Id.*, 11:27-12:2.

158. *Dolan* illustrates this type of Short Message in figure 5, which I have reproduced and annotated below:



Ex. 1005, FIG. 5 (annotated).

159. As *Dolan* explains, “[t]he Total Segments and Number of Enclosed Segment fields provide the means to identify and reassemble the segments of the coded audio information into the correct order at the end user device.” *Id.*, 4:5-7; *see also id.*, 10:11-14 (“The Total Segments field 524 and the Number of Enclosed Segment field 528 provide the means to identify and then correctly reassemble the segments of the coded audio information into the correct order at the end user device.”).

160. Therefore, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests breaking compressed audio information that is too large for one Short Message into multiple segments (“dividing the non-text representation into sections”) and placing each segment into a separate Short Message wherein each short Message is one of a total number of Short Messages (“and inserting each section into a body of a different corresponding text message of a plurality of text message”).

161. Additionally, in my opinion, a person having ordinary skill in the art would have understood that where compressed audio information is broken into multiple segments, the Short Message discussed with respect to claim 1 would be one of the plurality of Short Messages into which the multiple audio segments are placed (“wherein the first-mentioned text message is one of said plurality of text messages”).

- b. *11[b]: “further comprising sending the plurality of text messages over the wireless messaging channel from the mobile communication device to the recipient’s device.”*

162. In my opinion, *Dolan* discloses or suggests “sending the plurality of text messages over the wireless messaging channel from the mobile communication device to the recipient’s device.”

163. As I explained in Section X.A.1.e, *Dolan* discloses or suggests sending compressed audio information formatted into an SMS message (“sending the text message including the at least some of the non-text representation”) using an SMS protocol incorporated into at least one of the GSM, TDMA or CDMA wireless communication standards (“over the wireless messaging channel”) from one digital mobile telephone handset (“from the mobile communication device”) to a second digital mobile telephone handset (“to a recipient’s device”).

164. In my opinion, a person having ordinary skill in the art would have understood *Dolan*’s disclosure to equally apply to circumstances in which compressed audio information is formatted into multiple SMS messages.

165. More particularly, *Dolan* discloses with respect to figure 3 that “[o]nce one ***or more*** Short Messages are created to contain the compressed audio information, the End User Short Message Delivery Function – Transmit 315 initiates transmission of the Short Message(s) to the Message Center 106

where they are received by the Message Center Short Message Delivery Function – Receive 325.” Ex. 1005, 15:28-16:3; *see also id.* 11:22-25 (“[I]f the compressed audio information is too long to be contained in a single Short Message, it can be segmented into multiple Short Messages utilizing any common methodology known in the art. For example, SMS has a segmentation and reassembly feature already built in.”).

166. Therefore, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses or suggests that when the compressed audio information is formatted into multiple SMS messages, each of the plurality of SMS messages is sent (“sending the plurality of text messages”) using the SMS protocol incorporated into at least one of the GSM, TDMA or CDMA wireless communication standards (“over the wireless messaging channel”) from one digital mobile telephone handset (“from the mobile communication device”) to a second digital mobile telephone handset (“to the recipient’s device”).

B. Ground 1B: *Dolan* in view of *Freedman* Renders Obvious Claims 1, 4-7

167. In Ground 1B, I analyze the combination of *Dolan* and *Freeman*, in which a phoneme-based technique for compressing and decompressing audio as disclosed in *Freedman* is used to perform *Dolan*’s audio compression and decompression functions.

1. Claim 1

- a. ***1[pre] & 1[a]: 1[pre] (“A method of sending a voice message via a mobile communication device, said method comprising:”), and 1[a] (“receiving an utterance from a user of the mobile communication device;”)***

168. In my opinion, *Dolan* discloses or suggests 1[pre] and 1[a] for the same reasons I explained in Sections X.A.1.a and X.A.1.b, respectively.

- b. ***1[b]: “generating a non-text representation of the received utterance;”***

169. In my opinion, *Dolan* in combination with *Freedman* discloses or suggests “generating a non-text representation of the received utterance.”

170. As I already discussed in Section X.A.1.c, *Dolan* discloses or suggests compressing the audio message before inserting it into a Short Message, such as a SMS message. In Ground 1A, I focused on *Dolan*’s disclosure of vocoders that are already utilized by mobile phones to compress the audio message, but in fact, *Dolan* says that the audio information can be encoded by any of the many techniques known in the art. Ex. 1005, 7:10-16. Specifically, *Dolan* states:

[T]he captured information is encoded by means of any audio compression technology known in the art, such as the technology used in voice coding schemes on wireless networks, and the compressed information is stored at a message center or other facility that provides mailbox services to end users. ***Many***

techniques for compressing and encoding voice signals for digital transmission are well-known in the art, and it is contemplated that any of these techniques may be used to perform the audio compression function of the invention.

Id.

However, *Dolan* does not limit compression to vocoders. In fact, *Dolan* broadly discloses that “the captured [audio] information [can be] encoded by means of any audio compression technology known in the art” and that “[m]any techniques for compressing and encoding voice signals for digital transmission are well-known in the art. *Id.*, 7:10-14. *Dolan* is open to the idea “that any of these [known] techniques may be used to perform the audio compression function of the invention.” *Id.*, 7:15-16.

171. *Freedman* is an example of a known technique in the art for compressing audio information, and specifically audio comprising human voice/speed, using a phoneme-based encoding technique:

A speech-to-text-to-speech for use with on-line and real time transmission of speech with a small bandwidth from a source to a destination. *A speech is received and broken down to phonemes, which are encoded into series of symbols compatible with communication systems and other than a known symbolic representation of the speech in a known language for being transmitted through communication networks.* When received, the series of symbols is decoded to restore the phonemes and for

reconstituting a speech according to the phonemes prior to being communicated to a listening party.

Ex. 1006, Abstract.

172. *Freedman* provides “a system that allows for on-line transmission of speech with a small bandwidth from a source to a destination.” *Id.* ¶[0008]. To achieve this goal, *Freedman* breaks spoken words into phonemes, thereby translating the provided speech into a text that corresponds to the speech in a phonetic form. The phonetic text is then transmitted via a telecommunications network such as the Internet or a public telephony switching system to another device which restores the speech according to the phonemes received and provides the restored speech to a speaker. Specifically, *Freedman* states:

Referring to FIG. 2, a schematic representation of a speech-to-text-to-speech system is shown. Speech is provided to a processor 20 through an input device in the form, for example, of a microphone 21. The processor 20 breaks the spoken words down into *phonemes* and translates the provided speech *into a text that corresponds to the speech in a phonetic form*. The phonetic text is *transmitted via a telecommunication network such as the Internet or a public telephony switching system* and is *received by device 21* including a processor for *restoring the speech according to the phonemes received*. *The restored speech is then provided to an output port 22* in the form for example of a loud-speaker.

Id. ¶[0035].

173. As *Freedman* discloses, it translates speech into “phonetic form.”

Id. To generate this phonetic text, *Freedman* identifies the phonemes in received speech, transcribe the phonemes into a corresponding series of written characters, and encode those phonetic characters into a “symbolic representation,” stating specifically:

Referring to FIG. 3a, the speech to text portion of the system is in the form of a device including an input port 31 for receiving a speech to transform, the input port 31 for connecting with the microphone 21 for example. The input port 31 is in communication with a translator 32. ***The purpose of the translator is to identify the phonemes when a part of a speech is received at input port 31***, and to provide the identified phoneme to a text generator 33. ***The text generator*** is in the form for example of a phonetic word processor, which ***transcribes the phonemes into corresponding series of written characters such that a phonetic transcript of the original speech is generated***. The phonetic transcript is modified at 41 in order to render it compatible for telecommunication transmission system and communicated to the output port 34 and transmitted out. The modification includes ***encoding the phonetic characters***, the encoding resulting in a series of symbols as for example alphanumeric equivalent or ASCII codes according to look up tables. The encoding ***preferably results in a symbolic representation of the speech other than in a known human intelligible language***.

Id. ¶[0038].

Referring to FIG. 4a, the speech to text portion of the system is in the form of a device including an input port 42 for receiving speech to transform, the input port 42 for connecting with the microphone 21 for example. The input port 42 is in communication with a language selector 43 and optionally with a speaker identifier 44. The purpose of the language selector 43 is to identify a language of the speech prior to a communication session. Language identification is provided at the beginning of the transmitted phonetic text in the form of ENG for English, FRA for French, ITA for Italian and so forth. Upon identifying a language, a phoneme database is selected from the database 45 such that only the *phonemes corresponding to the identified language are used to translate the speech into a phonetic text by the text generator 46 in the form for example of a phonetic processor, which transcribes the phonemes into corresponding series of written characters such that a phonetic transcript of the original speech is generated.* When the system comprises a speaker identifier 44, the speaker provides his name or any indication of his identity, which is then associated with the generated text before being transmitted. The phonetic transcript is modified at 47 in order to render it compatible for telecommunication transmission system and communicated to the output port 48 and transmitted out. As will be apparent to those of skill in the art, *language dependent phonetic dictionaries allow for improved compression of the speech and for improved phoneme extraction.* Alternatively, language and regional characteristics are used to select a phonetic dictionary such as Sco

for Scottish English and Iri for Irish English in order to improve the phonetic dictionary for a particular accent or mode of speaking. Further alternatively, a speaker dependent phonetic dictionary is employed. Of course, it is preferable that a same dictionary is available at the receiving end for use in regenerating the speech in an approximately accurate fashion.

Id. ¶[0057].

174. *Freedman* provides an example of using its technique to encode the sentence: “hello, the sun is shining” as the following phonemes:

hə'ləʊðɪz sʌnɪzɪŋ aɪnɪŋ, which a POSITA would have understood to comprise a non-human intelligible string of symbols. Ex. 1006 ¶¶[0040]-[0041].

175. In my opinion, a person having ordinary skill in the art would have understood that *Freedman*'s phoneme-based technique for encoding speech would have resulted in a compressed form of the speech requiring less bandwidth for transmission than if the speech were transmitted without first being encoded as described by *Freedman*. For example, *Freedman* emphasizes the reduced or small bandwidth. Ex. 1006 ¶¶[0007] (“It would be advantageous to provide with a system that requires reduced bandwidth to support voice communication.”), [0008] (“[I]t is an object of the present invention to provide with a system that allows for on-line transmission of speech with a small bandwidth from a source to a destination.”).

176. *Freedman* discloses using phonemes to generate a “series of symbols” or “a symbolic representation of the speech other than in a known human intelligible language” representing speech. Ex. 1006 ¶[0038]. This is similar to the ’993 patent’s disclosure of using phonemes to generate a non-text representation of an utterance. *See* Ex. 1001, 2:33-39 (“Generating the non-text representation of the received utterance involves ... generat[ing] a string of symbols, wherein the string of symbols is the non-text representation. The symbols in the strig of symbols are selected from the group consisting of phonemes, diphones, and triphones (more specifically, the symbols are phonemes.”), 4:51-52 (“The output is a non-text representation of the spoken utterance, in this case, a phoneme string.”).

177. In my opinion, a person having ordinary skill in the art would have found it obvious to use *Freedman*’s phoneme-based technique to perform *Dolan*’s audio compression function. Indeed, *Dolan* explicitly provides the motivation to combine with other techniques for compressing/encoding voice signals, stating: “Many techniques for compressing and encoding voice signals for digital transmission are well-known in the art, and it is contemplated that any of these techniques may be used to perform the audio compression function of the invention.” Ex. 1005, 7:13-16.

178. In my opinion, a person having ordinary skill in the art would have found the combination obvious because it amounted to the simple substitution of one known element (*Freedman*'s phoneme-based technique to compress speech) for another (*Dolan*'s vocoder-based technique to compress speech) to obtain predictable results (to perform the speech compression function in *Dolan*'s system).

179. Further, in my opinion, a person having ordinary skill in the art would have had a reasonable expectation of success in combining *Dolan* and *Freedman* because both share a focus on enhancing communication efficiency through digital audio compression and transmission techniques. For example, a person having ordinary skill in the art would have understood that *Freedman*'s approach to minimizing bandwidth use with phonetic speech-to-text-to-speech conversion complements *Dolan*'s method of delivering audio via discrete messaging services. Moreover, in my opinion, this combination of *Dolan* and *Freedman* suggests a reasonable and acceptable solution for optimizing audio communication for devices with limited bandwidth, offering a solution that combines the benefits of short messaging and compression using phoneme-based encoding.

180. Accordingly, in my opinion, a person having ordinary skill in the art would have understood that the *Dolan-Freedman* combination discloses or

suggests generating a phoneme-based representation (“generating a non-text representation”) of the user’s utterance (“of the received utterance”).

- c. ***1[c] & 1[d]: 1[c] (“inserting at least some of the non-text representation into a body of a text message; and”), and 1[d] (“sending the text message including the at least some of the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device.”)***

181. In my opinion, *Dolan* in combination with *Freedman* discloses or suggests “inserting at least some of the non-text representation into a body of a text message” and “sending the text message including the at least some of the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device.”

182. *Dolan* discloses or suggests inserting the non-text representation of a user’s utterance (e.g., “Call home, emergency!”) into the body of a text message, and sending the text message including the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device for at least the reasons I discussed in Sections X.A.1.d and X.A.1.e, respectively. In the *Dolan-Freedman* combination, however, the non-text representation of the user’s utterance that is inserted into the body of the text message is a phonetic representation of the utterance. In my opinion, making this combination would have been obvious to a person having ordinary skill in the art for the reasons I discussed in Section X.B.1.b.

2. **Claim 4: “The method of claim 1, wherein generating the non-text representation of the received utterance involves performing recognition on a signal derived from the received utterance to generate a string of symbols, wherein the string of symbols is the non-text representation.”**

183. In my opinion, *Dolan* in combination with *Freedman* discloses or suggests “wherein generating the non-text representation of the received utterance involves performing recognition on a signal derived from the received utterance to generate a string of symbols, wherein the string of symbols is the non-text representation.”

184. As I previously described in Section X.B.1.b, in the *Dolan-Freedman* combination, *Freedman* discloses identifying the phonemes when a part of a speech is received (“performing recognition on a signal derived from the received utterance”) and encoding the phonetic characters preferably resulting in a symbolic representation of the speech other than in a known human intelligible language (“to generate a string of symbols, wherein the string of symbols is the non-text representation”).

3. **Claim 5: “The method of claim 4, wherein the symbols in the string of symbols are selected from the group consisting of phonemes, diphones, triphones, syllables, and demisyllables.”**

185. In my opinion, *Dolan* in combination with *Freedman* discloses or suggests “wherein the symbols in the string of symbols are selected from the

group consisting of phonemes, diphones, triphones, syllables, and demisyllables.”

186. I understand “group consisting of phonemes, diphones, triphones, syllables, and demisyllables” to be a Markush group. As such, I understand the prior art need disclose or suggest only one item from the group in order to meet this feature.

187. *Freedman* discloses “identify[ing] the phonemes when a part of a speech is received ..., [and] provid[ing] the identified phoneme to a text generator ... [to] transcribe[] the phonemes into corresponding series of written characters such that a phonetic transcript of the original speech is generated ... [and] encoding the phonetic characters, the encoding resulting in a series of symbols as for example alphanumeric equivalent or ASCII codes according to look-up tables. The encoding preferably result[ing] in a symbolic representation of the speech other than in a known human intelligible language.” Ex. 1006, ¶[0038].

188. Accordingly, in my opinion, a person having ordinary skill in the art would have understood that the *Dolan-Freedman* combination discloses or suggests the symbolic representation of the user’s utterance (“the symbols in the string of symbols”) being based on phonemes identified in the utterance (“are selected from the group consisting of phonemes . . .”).

4. **Claim 6: The method of claim 4, wherein the recognition is phoneme recognition and the symbols in the generated string of symbols are phonemes.**

189. In my opinion, *Dolan* in combination with *Freedman* discloses or suggests this claim for the reasons I discussed in Sections X.B.2, X.B.3, as well as X.B.1.b.

5. **Claim 7: “The method of claim 1, wherein generating the non-text representation of the received utterance involves performing phoneme recognition on a signal derived from the received utterance to generate a phoneme string, wherein the phoneme string is the non-text representation.”**

190. In my opinion, *Dolan* in combination with *Freedman* discloses or suggests this claim for the reasons I discussed in Sections X.B.2, X.B.3, as well as X.B.1.b.

C. Ground 2A: *Dolan* in view of *Halonen* Renders Obvious Claims 12, 15-17

191. Ground 2A is directed to a combination of *Dolan* and *Halonen*, in which the hardware from a typical cellular phone as in *Halonen* is used to perform *Dolan*’s transmission and reception of audio information over SMS.

1. Claim 12

- a. ***12[pre]: “A mobile communication device for sending a voice message, said mobile communication device comprising:”***

192. I understand that “[a] method of sending a voice message via a mobile communication device, said method comprising” is the preamble to

claim 1 of the '078 patent. I have been asked to assume that the preamble is a claim limitation. Under that assumption, in my opinion, *Dolan* discloses or suggests a digital mobile telephone handset (“a mobile communication device”) for sending an audio message (“for sending a voice message”) for the reasons I discussed in Section X.A.1.a.

- b. 12[a] through 12[d]: 12[a] (“a processor system;”), 12[b] (“a microphone for receiving an utterance from a user of the mobile communication device;”), 12[c] (“a transceiver; and”), 12[d] (“memory storing code which when executed on the processor system causes the mobile communication device to:”)***

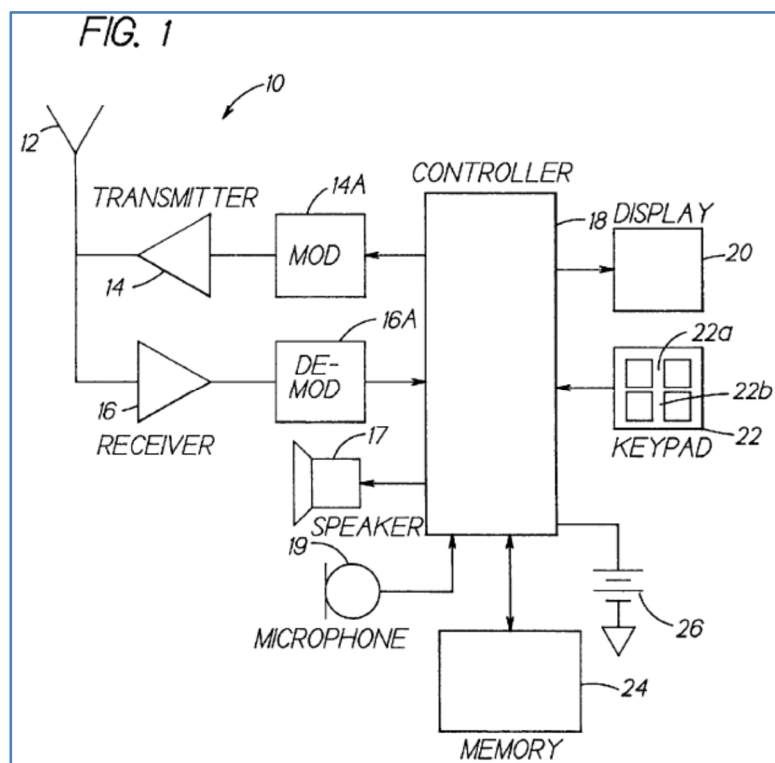
193. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests these features.

194. For the reasons I discussed in Section X.A.2, a person having ordinary skill in the art would have understood *Dolan*’s digital mobile telephone handset to be a cellular phone.

195. In my opinion, a person having ordinary skill in the art would have been motivated to look for additional information regarding how to implement *Dolan*’s digital mobile telephone handset (cellular phone) because *Dolan* itself does not focus on the implementation details of its digital mobile telephone handset. In searching for such a reference, in my opinion, a person having ordinary skill in the art would have readily identified *Halonen*.

196. *Halonen* discloses the hardware components that are incorporated into a typical cellular phone. Ex. 1007, 2:43-45 (“Reference is made to FIGS. 1 and 2 for illustrating a wireless user or mobile terminal 10, such as but not limited to a cellular radiotelephone.”). Similar to *Dolan*’s digital mobile telephone handset, *Halonen*’s mobile terminal implements one or more standard cellular communication protocols. *Id.*, 3:18-25 (“It should be understood that the mobile terminal 10 ... may be capable of operating with any of a number of air interface standards, including but not limited ... GSM ... and IS-95 (CDMA).”). Another similarity between *Dolan*’s digital mobile telephone handset and *Halonen*’s mobile terminal is that both can process SMS messages. *Id.*, 4:55-57 (“By example, the code blocks can be sent using so-called Short Message Service (SMS) messages.”).

197. *Halonen* discloses its mobile terminal 10 has a “controller 18 ... comprised of a digital signal processor (DSP) 18a and a microprocessor control unit (MCU) 18b” (*id.*, 3:38-40), a microphone 19 (2:61), and “a transceiver (14, 16) for conducting bidirectional RF communications with a communications network (32)” (*id.*, Abstract). *See also id.*, 1:63-64, 2:8-9, claim 1 (“A wireless user terminal comprising a transceiver for conducting bidirectional RF communications with a communications network.”), claims 3, 4, 14, 16, FIG. 1.



Ex. 1007, FIG. 1.

198. By repeatedly stating that bidirectional RF communication is conducted by a “transceiver,” a person having ordinary skill in the art would have understood *Halonen*’s disclosure to encompass using a single component to implement *Halonen*’s transmitting and receiving functions even though *Halonen*’s figure 1 separately labels the transmitter and receiver functions of such a transceiver. *See* Ex. 1007, 1:63-64, 2:8-9, claims 1, 3, 4, 14, 16. Moreover, transceivers were well-known and widely used in mobile/cellular phones by April 2004. For example, *Kimball*, like *Halonen*, is directed to a cellular telephone that uses the CDMA standard (Ex. 1012, 2:53-58), a microphone/speaker, (*id.*, 5:44-47, 6:47-50), memory storing code executed by

the processor, (*id.*, 6:51-60), and a transceiver, (*id.*, 5:40-47, FIGs. 3-5). As depicted in *Kimball*'s figure 3, the transceiver is a single component.⁸

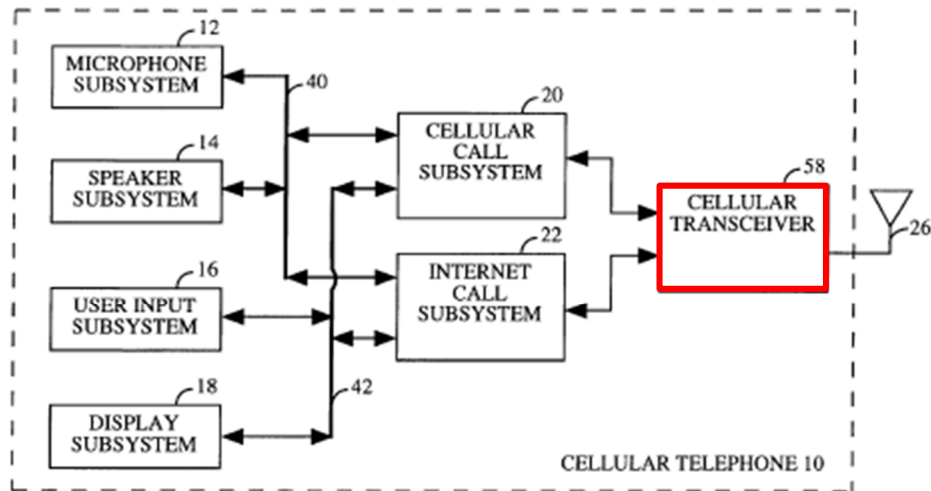


FIG. 3

Ex. 1012, FIG. 3 (annotated).

Thus, even if one were to assume that *Halonen* does not disclose implementing its transceiver as a single component (which, as I explain above, it does), a person having ordinary skill in the art would have found it obvious to implement *Halonen*'s transmitting and receiving functions using a transceiver (instead of, for example, separate transmitter and receiver) because transceivers were well

⁸ To be clear, here I am relying on Ex. 1012 (*Kimball*) as evidence of a person having ordinary skill in the art's knowledge and the state of the art. *Kimball* is not part of the unpatentability grounds.

known in the art and doing so would have amounted to substituting one known element (a transceiver) for another (separate transmitter and receiver) to obtain predictable results (to effectuate transmitting and receiving).

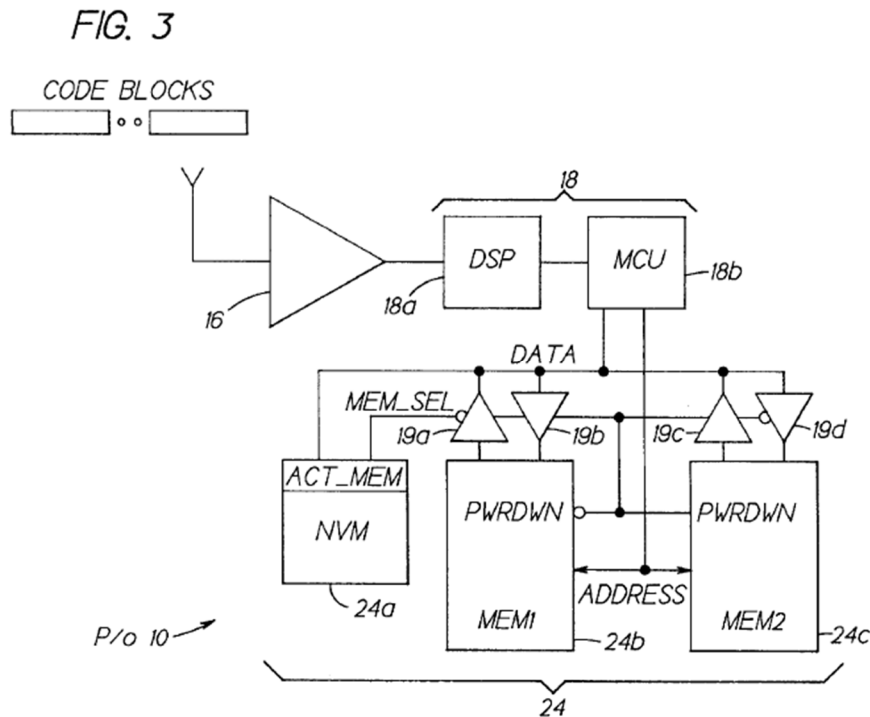
199. In my opinion, a person having ordinary skill in the art would have understood that microphone 19 of *Halonen*'s mobile terminal 10 would have been configured to and been suitable for receiving words spoken by the user of the mobile terminal, such as, "Call home, emergency!", and converting the sounds spoken by the user to electrical signals for further processing by the mobile terminal 10.

200. In addition to the other components, *Halonen* also discloses that its mobile terminal has memory storing code to be executed by the processor. Specifically, *Halonen* discloses:

The mobile terminal 10 also includes *various memories, shown collectively as the memory 24*, wherein are stored a plurality of constants and variables that are used by the controller 18 during the operation of the mobile terminal. For example, the memory 24 stores the values of various cellular system parameters and the number assignment module (NAM). *An operating program for controlling the operation of controller 18 is also stored in the memory 24 (typically in a ROM device).*

Ex. 1007, 3:1-9; *see also id.*, 3:50-52 ("The software routines executed by the MCU 18b to operate the mobile terminal 10 are referred to herein as its

operating program.”), 3:64-66 (“The memories 24b and 24c are provided to store the operating program for the MCU 18b.”), FIG. 3.



Ex. 1007, FIG. 3.

201. In my opinion, a person having ordinary skill in the art would have understood that “code” in the context of the ’078 patent refers to “program instructions” and more particularly, here, to “[m]achine code consist[ing] of numerical instruction that the computer can recognize and execute and that were converted from source code.” See Ex. 1013 at 91 (defining “code” as “Program instructions. Source code consists of human-readable statements written by a programmer in a programming language. Machine code consists of numerical instructions that the computer can recognize and execute and that were

converted from source code.”). Therefore, a person having ordinary skill in the art would have understood that the operating program stored in *Halonen*’s memory 24 comprises program instructions (i.e., code) that when executed by *Halonen*’s controller 18 (e.g., DSP 18a and/or MCU 18b) causes the controller to perform the programmed functionality.

202. In my opinion, a person having ordinary skill in the art would have appreciated that the mobile communications devices in *Dolan* and *Halonen* are similar to each other in as much as both are what are commonly referred to as “cellular phones” and both are compatible with and configured to implement one or more common cellular telephone protocols such as GSM and/or CDMA. *See* Section X.A.1.e (discussing why a person having ordinary skill in the art would have understood *Dolan*’s end user device to be a cellular phone that transmits over a wireless channel and incorporates protocols such as TDMA, GSM and/or CDMA).

203. In my opinion, a person having ordinary skill in the art would have understood that *Dolan*’s digital mobile telephone handset would have had these same components as *Halonen*. However, to the extent that *Dolan* does not expressly disclose the components of its digital mobile telephone handset, in my opinion, a person having ordinary skill in the art looking for additional implementation details would have readily identified *Halonen* and would have

understood the applicability *Halonen*'s teaching of a controller, a microphone, transceiver, and a memory to *Dolan*'s digital mobile telephone handset.

204. Further, based on the similar cellular telephone-type devices of *Dolan* and *Halonen*, in my opinion, a person having ordinary skill in the art would have understood how to predictably apply *Halonen*'s teaching regarding the components of a cellular telephone to *Dolan*'s digital mobile telephone handset.

205. Thus, in my opinion, a person having ordinary skill in the art would have considered the combination to amount to the use of a known technique (*Halonen*'s teachings regarding the components of a cellular telephone, including a controller, a microphone, a transceiver, and a memory) to improve similar devices (*Dolan* digital mobile telephone handset) in the same way (to incorporate components into *Dolan*'s digital mobile telephone handset to implement such a device).

206. Accordingly, in my opinion, a person having ordinary skill in the art would have understood that the *Dolan-Halonen* combination discloses or suggests a digital mobile telephone handset ("mobile communication device") having a controller comprising a DSP and a MCU ("processor system"), a microphone for receiving a an audio message such as "Call home, emergency!" spoken by a user ("a microphone for receiving an utterance from a user of the

mobile communication device”), a transceiver (“transceiver”), and a memory for storing an operating program for controlling the operation of the mobile terminal (“memory storing code which when executed on the processor system causes the mobile device to”).

207. Additionally, *Halonen* discloses that the operating program controls the operation of the mobile terminal 10.Ex. 1007, 3:7-9, 3:50-52, 3:64-66. As such, in my opinion the *Dolan-Halonen* combination discloses or suggests that when the operating program stored in the memory is executed by the controller, the controller causes the digital mobile telephone handset to perform operations that meet claim elements 12[e] through 12[g] and the dependent claims as discussed below.

c. 12[e]: “generate a non-text representation of the received utterance;”

208. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests “generate a non-text representation of the received utterance.”

209. More particularly, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses its digital mobile telephone handset performing this functionality for the reasons I discussed in Section X.A.1.c. As I already discussed in Section X.C.1.b, in the *Dolan-Halonen* combination, when the operating program stored in the memory is executed by

the controller, the controller causes the digital mobile telephone handset to perform this operation.

d. 12[f]: “insert the non-text representation into a body of a text message; and”

210. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests “insert the non-text representation into a body of a text message.”

211. More particularly, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* discloses its digital mobile telephone handset performing this functionality for the reasons I discussed in Section X.A.1.d. As I noted in Section X.A.1.d, *Dolan*’s figure 4 discloses an embodiment in which all of the compressed audio information is inserted into one SMS message.

212. Additionally, as I discussed in Section X.C.1.b, in the *Dolan-Halonen* combination, when the operating program stored in the memory is executed by the controller, the controller causes the digital mobile telephone handset to perform this operation.

e. 12[g]: “send the text message including the non-text representation via the transceiver over a wireless messaging channel from the mobile communication device to a recipient’s device.”

213. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests “send the text message including the non-text representation via the

transceiver over a wireless messaging channel from the mobile communication device to a recipient's device.”

214. More particularly, *Dolan* discloses its digital mobile telephone handset “send[ing] the text message including the non-text representation . . . over a wireless messaging channel from the mobile communication device to a recipient's device” for the reasons I discussed in Section X.A.1.e.

215. In the *Dolan-Halonen* combination, in my opinion, it would have been obvious to a person having ordinary skill in the art to send the text message via the transceiver because *Halonen* teaches that the transceiver is used to bidirectionally communicate with other devices on a network. Specifically, *Halonen* discloses:

A wireless user terminal (10), such as a cellular telephone, includes ***a transceiver (14, 16) for conducting bidirectional RF communications with a communications network (32)*** and a data processor (18b) for controlling the operation of the terminal. The terminal further includes (a) a first memory capable of storing a data processor program; (b) a second memory capable of storing a data processor program; (c) a memory device or register for storing a flag for indicating which of the first and second memories is a currently Active memory, that stores a currently active data processor program, and which of the first and second memories is a currently Idle memory; and (d) circuitry for receiving another data processor program from the communications network and for

storing the received another data processor program into the Idle memory. The circuitry operates while the data processor executes the data processor program stored in the Active memory.

Ex. 1007, Abstract.

1. *A method for operating a wireless user terminal having a transceiver for conducting bidirectional RF communications with a communications network*, comprising the steps of: operating the terminal with a first computer program that is stored in a first memory; the step of operating including a step of receiving a second computer program from the communications network and storing the received second computer program into a second memory; and after the second computer program is stored in the second memory, operating the terminal with the second computer program; wherein the step of receiving includes the step of temporarily terminating the receipt of the second computer program in the event an operator of the mobile terminal requires the use of the transceiver.

Id., claim 1; *see also id.* 1:63-64, 2:8-9, claims 3, 4, 14, 16, Fig. 1.

216. Additionally, as I already discussed in Section X.C.1.b, in the *Dolan-Halonen* combination, when the operating program stored in the memory is executed by the controller, the controller causes the digital mobile telephone handset to perform this operation.

2. **Claim 15: “The mobile communication device of claim 12, wherein the wireless messaging channel is an SMS channel and the text message is an SMS message.”**

217. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests this claim for the reasons I discussed above in Section X.A.4.

3. **Claim 16: “The mobile communication device of claim 12, wherein the non-text representation is a compressed version of the received utterance.”**

218. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests this claim for the reasons I discussed in Section X.A.6.

4. **Claim 17**

- a. ***Claim 17[a]: “The mobile communication device of claim 12, wherein the code when executed on the processor further causes the mobile communication device to: divide the non-text representation into sections; insert each section into a body of a different corresponding text message of a plurality of text messages, wherein the first-mentioned text message is one of said plurality of text messages; and”***

219. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests “wherein the code when executed on the processor further causes the mobile communication device to: divide the non-text representation into sections; insert each section into a body of a different corresponding text message of a plurality of text messages, wherein the first-mentioned text message is one of said plurality of text messages.”

220. More particularly, in my opinion a person having ordinary skill in the art would have understood that *Dolan* discloses its digital mobile telephone handset performing these functionalities for the reasons I discussed in Section X.A.7.a.

221. As I discussed above in Section X.C.1.b, in the *Dolan-Halonen* combination, when the operating program stored in the memory is executed by the controller, the controller causes the digital mobile telephone handset to perform this operation.

b. 17[b]: “send the plurality of text messages over the wireless messaging channel from the mobile communication device to the recipient’s device.”

222. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests “send the plurality of text messages over the wireless messaging channel from the mobile communication device to the recipient’s device.”

223. More particularly, in my opinion a person having ordinary skill in the art would have understood that *Dolan* discloses its digital mobile telephone handset performing these functionalities for the reasons I discussed in Section X.A.7.b.

224. As I previously discussed in Section X.C.1.b, in the *Dolan-Halonen* combination, when the operating program stored in the memory is executed by

the controller, the controller causes the digital mobile telephone handset to perform this operation.

D. Ground 2B: *Dolan* in view of *Halonen* and *Freedman* Renders Obvious Claims 12, 15-17

225. In this Ground 2B, I provide my opinions regarding a combination of *Dolan*, *Halonen*, and *Freedman*, in which a phoneme-based technique for compressing audio as described in *Freedman* is used to perform *Dolan*'s audio compression function, and *Halonen* provides additional details regarding the components incorporated into *Dolan*'s digital mobile telephone handset. A POSITA would have been motivated to combine *Dolan*, *Halonen*, and *Freedman* for the same reasons I discussed in Section X.B.1.b (*Dolan* and *Freedman*), and Section X.C.1.b (*Dolan* and *Halonen*).

1. Claim 12

- a. *12[pre] through 12[d]: 12[pre] ("A mobile communication device for sending a voice message, said mobile communication device comprising:"), 12[a] ("a processor system;"), 12[b] ("a microphone for receiving an utterance from a user of the mobile communication device;"), 12[c] ("a transceiver; and"), 12[d] ("memory storing code which when executed on the processor system causes the mobile communication device to:")*

226. In my opinion, *Dolan* in combination with *Halonen* discloses or suggests these features for the reasons I discussed in Sections X.C.1.a and X.C.1.b.

b. 12[e]: “generate a non-text representation of the received utterance;”

227. In my opinion, *Dolan* in combination with *Halonen* and *Freedman* discloses or suggests “generate a non-text representation of the received utterance.”

228. More particularly, in my opinion a person having ordinary skill in the art would have understood that *Dolan* in combination with *Freedman* discloses or suggests generating a phoneme-based representation (‘generating a non-text representation’) of the user’s utterance (“of the received utterance”) for the reasons I discussed in Section X.B.1.b.

229. As I also discussed in Section X.C.1.b, *Dolan* in combination with *Halonen* discloses or suggests that, when the operating program stored in the memory is executed by the controller, the controller causes the digital mobile telephone handset to perform this operation.

230. In my opinion, *Freedman* and *Halonen* are compatible with each other because *Freedman* discloses that its phoneme-based speech processing is performed using a processor 20. Ex. 1006, ¶[0035]. In my opinion, a person having ordinary skill in the art would have understood, therefore, that in the *Dolan-Halonen-Freedman* combination, *Freedman*’s teachings regarding how to encode speech using phonemes would be implemented through the operating

program stored in the memory of, and executed by the controller of, the *Dolan-Halonen* combination.

- c. ***12[f] & 12[g]: 12[f] (“insert the non-text representation into a body of a text message; and”), and 12[g] (“send the text message including the non-text representation via the transceiver over a wireless messaging channel from the mobile communication device to a recipient’s device.”)***

231. In my opinion, *Dolan* in combination with *Halonen* and *Freedman* discloses or suggests “insert the non-text representation into a body of a text message; and,” and “send the text message including the non-text representation via the transceiver over a wireless messaging channel from the mobile communication device to a recipient’s device.”

232. More particularly, in my opinion, a person having ordinary skill in the art would have understood that *Dolan* in combination with *Freedman* discloses or suggests inserting a phonetic representation of the user’s utterance (“insert the non-text representation”) into the body of a text message, and sending the text message including the non-text representation over a wireless messaging channel from the mobile communication device to a recipient’s device for the reasons I discussed in Section X.B.1.c. As I noted in Section X.A.1.d, with reference to figure 4, *Dolan* discloses an embodiment in which all of the compressed audio message is inserted into one SMS message.

233. In the *Dolan-Halonen-Freedman* combination, in my opinion it would have been obvious to a person having ordinary skill in the art to send the text message via the transceiver for the reasons discuss in Section X.C.1.e. Additionally, for the reasons I discussed in Section X.C.1.b, in the *Dolan-Halonen-Freedman* combination, when the operating program stored in the memory is executed by the controller, the controller causes the digital mobile telephone handset to perform this operation.

2. Claim 13: “The mobile communication device of claim 12, wherein generating the non-text representation of the received utterance involves performing recognition on a signal derived from the received utterance to generate a string of symbols, the string of symbols being the non-text representation.”

234. In my opinion, *Dolan* in combination with *Halonen* and *Freedman* discloses or suggests “wherein generating the non-text representation of the received utterance involves performing recognition on a signal derived from the received utterance to generate a string of symbols, the string of symbols being the non-text representation.”

235. Specifically, in my opinion, a person having ordinary skill in the art would have understood that in the *Dolan-Halonen-Freedman* combination, *Freedman* discloses identifying the phonemes when a part of a speech is received (“performing recognition on a signal derived from the received utterance”) and encoding the phonetic characters preferably resulting in a

symbolic representation of the speech other than in a known human intelligible language (“to generate a string of symbols, wherein the string of symbols is the non-text representation”) for the reasons I discussed in Section X.B.2, as well as Section X.B.1.b.

3. Claim 14: “The mobile communication device of claim 13, wherein the symbols in the string of symbols are selected from the group consisting of phonemes, diphones, triphones, syllables, and demisyllables.”

236. In my opinion, *Dolan* in combination with *Halonen* and *Freedman* discloses or suggests “wherein the symbols in the string of symbols are selected from the group consisting of phonemes, diphones, triphones, syllables, and demisyllables.”

237. More particularly, a person having ordinary skill in the art would have understood that *Freedman* discloses this claim for the reasons I discussed in Section X.B.3.

XI. SECONDARY CONSIDERATIONS

238. I am not aware of any evidence in the '078 patent's prosecution history or elsewhere supporting any secondary considerations arguments, or evidence of nexus of such alleged evidence to the Challenged Claims. *See generally* Ex. 1004. To the extent Patent Owner asserts the existence of any

secondary considerations in its responses, I reserve the right to address any such evidence.

XII. CONCLUSION

I declare that all statements made herein of my knowledge are true, that all statements made on information and belief are believed to be true, and that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Executed on: July 22, 2024

Mark R. Lanning
Mark R. Lanning