

Filed: August 31, 2023

Filed on behalf of Meta Platforms, Inc.

By: Lisa Nguyen
David Tennant
Eric E. Lancaster
Daniel Margolis
Grace Wang
Alan Billharz
Colby Davis

Allen & Overy LLP
550 High Street
Palo Alto, CA 94301
Telephone: 650-388-1724
Facsimile: 202-683-3999
Email: AO_Meta_Jawbone@AllenOvery.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

META PLATFORMS, INC.,
Petitioner,

v.

JAWBONE INNOVATIONS, LLC,
Patent Owner.

IPR2023-01352
U.S. Patent No. 7,246,058

**PETITION FOR *INTER PARTES* REVIEW
OF CLAIMS 1-5 OF U.S. PATENT NO. 7,246,058**

TABLE OF CONTENTS

I.	Precise Relief Requested	1
II.	Background.....	3
	A. The '058 Patent	3
	B. Prosecution History	6
III.	Level of Ordinary Skill in the Art	7
IV.	Claim Construction.....	7
V.	The Challenged Claims Are Unpatentable and Should Be Cancelled	8
	A. [Ground 1] Claims 1-2 and 4-5 Are Rendered Obvious by Meisel in View of Kaneda and Freidin	8
	1. Meisel.....	8
	2. Kaneda	13
	3. Freidin	15
	4. Independent Claim 1	17
	a. [1Preamble] “A system for detecting voiced and unvoiced speech in acoustic signals having varying levels of background noise, comprising:”.....	17
	b. [1a] “at least two microphones that receive the acoustic signals;”	20
	c. [1b] “at least one voicing sensor that receives physiological information associated with human voicing activity; and”.....	21
	d. [1c] “at least one processor coupled among the microphones and the voicing sensor,”	23
	e. [1d] “wherein the at least one processor; generates cross correlation data between the physiological information	

	and an acoustic signal received at one of the two microphones;”	26
i.	Motivation to combine	32
ii.	Expectation of success	33
f.	[1e] “identifies information of the acoustic signals as voiced speech when the cross correlation data corresponding to a portion of the acoustic signal received at the one receiver exceeds a correlation threshold;”	37
i.	Motivation to combine	40
ii.	Expectation of success	42
g.	[1f] “generates difference parameters between the acoustic signals received at each of the two receivers, wherein the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals;”	43
h.	[1g] “identifies information of the acoustic signals as unvoiced speech when the difference parameters exceed a gain threshold; and”	50
i.	[1h] “identifies information of the acoustic signals as noise when the difference parameters are less than the gain threshold.”	56
i.	Motivation to combine	59
ii.	Expectation of success	63
5.	Independent Claim 2	65
a.	[2Preamble] “A method for removing noise from acoustic signals, comprising:”	65
b.	[2a] “receiving the acoustic signals at two receivers and receiving physiological information associated with human voicing activity at a voicing sensor;”	66

c.	[2b] “generating cross correlation data between the physiological information and an acoustic signal received at one of the two receivers;”	66
d.	[2c] “identifying information of the acoustic signals as voiced speech when the cross correlation data corresponding to a portion of the acoustic signal received at the one receiver exceeds a correlation threshold;”	66
e.	[2d] “generating difference parameters between the acoustic signals received at each of the two receivers, wherein the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals;”	67
f.	[2e] “identifying information of the acoustic signals as unvoiced speech when the difference parameters exceed a gain threshold; and”	67
g.	[2f] “identifying information of the acoustic signals as noise when the difference parameters are less than the gain threshold.”	67
6.	Claim 4: “further comprising performing denoising on the identified noise.”	67
7.	Claim 5: “wherein the voicing sensor includes at least one detector selected from a group including radio frequency devices, electroglottographs, ultrasound devices, acoustic throat microphones, and airflow detectors.”	68
B.	[Ground 2] Claim 3 Is Rendered Obvious by Meisel in View of Kaneda, Freidin, and Lyons	69
1.	Lyons.....	69
2.	Claim 3: “further comprising generating the gain threshold using standard deviations corresponding to the generation of the difference parameters.”	70
i.	Motivation to combine	72

ii.	Expectation of success	73
C.	[Ground 3] Claim 4 Is Rendered Obvious by Meisel in View of Kaneda, Freidin, and Sakoe.....	74
i.	Motivation to combine	74
ii.	Expectation of success	75
VI.	DISCRETIONARY DENIAL UNDER §314(A) IS NOT APPROPRIATE.	76
A.	Co-Pending Litigation Does Not Warrant Discretionary Denial (<i>Fintiv</i>)	76
1.	Factor 1: Potential Stay	76
2.	Factor 2: Proximity of Trial to FWD	77
3.	Factor 3: Investment in Parallel Proceeding	78
4.	Factor 4: Overlapping Issues	78
5.	Factor 5: The Parties	78
6.	Factor 6: Other Circumstances.....	79
B.	Existence of Previously Filed IPRs Does Not Warrant Discretionary Denial (<i>General Plastic</i>)	79
VII.	DISCRETIONARY DENIAL UNDER §325(D) IS NOT APPROPRIATE.	81
VIII.	Mandatory Notices.....	82
A.	Real Parties-in-Interest Under 37 C.F.R. § 42.8(b)(1).....	82
B.	Related Matters Under 37 C.F.R. § 42.8(b)(2)	82
C.	Lead and Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3)	85
D.	Service Information Under 37 C.F.R. § 42.8(b)(4).....	87
IX.	Grounds for Standing.....	87

X.	Conclusion	87
----	------------------	----

TABLE OF AUTHORITIES

	Page(s)
 Cases	
<i>Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH,</i> IPR2019-01469, Paper 6 (Feb. 13, 2020)	79
<i>Apple v. Fintiv,</i> IPR2020-00019, Paper 11 (P.T.A.B. Mar. 20, 2020)	73, 77
<i>GAF Materials LLC v. Kirsch Research and Dev., LLC,</i> IPR2021-00192, Paper 14 (P.T.A.B. May 25, 2021)	75
<i>General Plastic Indus. Co., Ltd. v. Canon Kabushiki Kaisha,</i> IPR2016-01357, Paper 19 (P.T.A.B. Sept. 6, 2017)	77, 78, 79
<i>Global Tel*Link Corp. v. HLFIP Holding, Inc.,</i> IPR2021-00444, Paper 14 (P.T.A.B. Jul. 22, 2021)	75
<i>In re Google LLC,</i> 58 F.4th 1379 (Fed. Cir. 2023)	74
<i>Google LLC v. Jawbone Innovations, LLC,</i> IPR2022-00630, Paper 10 (P.T.A.B. Sept. 13, 2022)	76, 77
<i>In re GPAC Inc.,</i> 57 F.3d 1573 (Fed. Cir. 1995)	7
<i>Huawei Techs. Co. v. WSOU Invs., LLC,</i> IPR2021-00226, Paper 10 (P.T.A.B. Jun. 10, 2021)	76
<i>Jawbone Innovations, LLC v. Amazon.com, Inc., et al.,</i> No. 2:21-cv-00435-JRG, Dkt. No. 60 (E.D. Tex. Sept. 21, 2022)	74, 82
<i>Jawbone Innovations, LLC v. Meta Platforms, Inc.,</i> 6:23-cv-00158 (W.D. Tex. Feb. 28, 2023)	2
<i>Jawbone Innovations, LLC v. Meta Platforms, Inc.,</i> No. 6:23-cv-00158, Dkt. No. 1 (W.D. Tex. Feb. 28, 2023)	74

<i>Mercedes-Benz USA, LLC v. Carucel Invs. L.P.</i> , IPR2019-01404, Paper 12 (P.T.A.B. Jan. 22, 2020)	78
<i>NetNut Ltd. v. Bright Data Ltd.</i> , IPR2021-00465, Paper 11 (P.T.A.B. Aug. 12, 2021).....	78
<i>Robert Bosch Healthcare Sys., Inc. v. Cardiocom, LLC</i> , No. C-14-1575 EMC, 2014 WL 3107447 (N.D. Cal. July 3, 2014)	74
<i>Samsung Elecs. Am. Inc. v. Snik LLC</i> , IPR2020-01428, Paper 10 (P.T.A.B. Mar. 9, 2021).....	76
<i>Sand Revolution II, LLC v. Cont’l Intermodal Grp.-Trucking LLC</i> , IPR2019-01393, Paper 24 (P.T.A.B. Jun. 16, 2020)	76
<i>Skechers U.S.A., Inc. v. Nike, Inc.</i> , IPR2021-00160, Paper 10 (P.T.A.B. May 19, 2021)	76
<i>Symantec Corp. v. Zscaler, Inc.</i> , No. 17-cv-04426-JST, 2018 WL 3539267 (N.D. Cal. July 23, 2018)	74
<i>Toshiba Am. Info. Sys., Inc. v. Walletex Microelecs. Ltd.</i> , IPR2018-01538, Paper 11 (P.T.A.B. Mar. 5, 2019).....	78, 79
<i>Twilio, Inc. v. TeleSign Corp.</i> , No. 16-cv-06925-LHK, 2018 WL 1609630 (N.D. Cal. April 3, 2018)	74
<i>Uniloc USA Inc. v. LG Elecs., U.S.A. Inc.</i> , No. 18-cv-06737-JST, 2019 WL 1905161 (N.D. Cal. April 29, 2019)	74
<i>Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.</i> , 200 F.3d 795 (Fed. Cir. 1999)	7

Statutes

35 U.S.C. § 102(a)	1, 2
35 U.S.C. § 102(b)	1, 2
35 U.S.C. § 102(e)	1
35 U.S.C. § 103	1, 2

35 U.S.C. § 311	1
35 U.S.C. § 325(d)	79, 80

Other Authorities

37 C.F.R. § 42.8(b)(1)	80
37 C.F.R. § 42.8(b)(2)	80
37 C.F.R. § 42.8(b)(3)	82
37 C.F.R. § 42.8(b)(4)	83
37 C.F.R. § 42.10(b)	83
37 C.F.R. § 42.24(a)	1
37 C.F.R. § 42.24(d)	1
37 C.F.R. § 42.100(b)	7

LIST OF EXHIBITS

Exhibit	Description
Ex. 1001	U.S. Patent No. 7,246,058 B2 (“the ’058 patent”)
Ex. 1002	File History of U.S. Patent No. 7,246,058
Ex. 1003	Declaration of Dr. David V. Anderson, Ph.D.
Ex. 1004	CV of Dr. David V. Anderson, Ph.D.
Ex. 1005	U.S. Patent No. 4,718,096 (“Meisel”)
Ex. 1006	U.S. Patent No. 5,208,864 (“Kaneda”)
Ex. 1007	U.S. Patent No. 6,148,313 (“Freidin”)
Ex. 1008	U.S. Patent No. 4,239,936 (“Sakoe”)
Ex. 1009	Signal Processing Toolbox for Use with MATLAB, User’s Guide, Version 4.2 (January 1999) (“MATLAB Manual”)
Ex. 1010	<i>Reserved</i>
Ex. 1011	Steven W. Smith, The Scientist and Engineer’s Guide to Digital Signal Processing (2d ed. 1999) (“Smith”)
Ex. 1012	Richard G. Lyons, <i>Understanding Digital Signal Processing</i> (1996) (“Lyons”)
Ex. 1013	U.S. Patent No. 5,724,485 (“Rainton”)
Ex. 1014	U.S. Patent No. 5,774,847 (“Chu”)
Ex. 1015	U.S. Patent No. 6,694,028 (“Matsuo”)
Ex. 1016	Vinay K. Ingle and John G. Proakis, <i>Digital Signal Processing Using MATLAB</i> (2000) (“Ingle”)
Ex. 1017	U.S. Patent No. 6,219,645 (“Byers”)
Ex. 1018	U.S. Patent No. 6,246,760 (“Makino”)
Ex. 1019	<i>Reserved</i>
Ex. 1020	U.S. Patent No. 6,760,449 (“Matsuo 2”)
Ex. 1021	U.S. Patent No. 6,766,019 (“Benesty”)
Ex. 1022	<i>Reserved</i>

Exhibit	Description
Ex. 1023	Richard O. Duda and Peter E. Hart, <i>Pattern Classification and Scene Analysis</i> (1973) (“Duda”)
Ex. 1024	David M. Green and John A. Swets, <i>Signal Detection Theory and Psychophysics</i> (1988) (“Green”)
Ex. 1025	J. R. Deller et al., “ShortTerm Processing of Speech,” in <i>Discrete-Time Processing of Speech Signals</i> , IEEE (2000) (“Deller”)
Ex. 1026	<i>Reserved</i>
Ex. 1027	Declaration of Lauren Gluckman regarding Lyons
Ex. 1028	Online Copyright Office Public Catalog Record of Lyons

I. Precise Relief Requested

Petitioner Meta Platforms, Inc. (“Petitioner” or “Meta”), requests review under 35 U.S.C. § 311 of claims 1-5 (“challenged claims”) of U.S. Patent No. 7,246,058 (“the ’058 patent”) and cancellation of the challenged claims under pre-AIA 35 U.S.C. § 103 based on the prior art references and grounds of unpatentability set forth below. The grounds presented in this petition are identical to the grounds previously presented to the Office in IPR2022-00623. IPR2022-00623 was instituted but was dismissed following institution and before oral argument due to settlement. Petitioner therefore asks that the Office reinstitute proceedings for the ’058 patent for the same reasons the Office previously instituted, as set forth below:

Prior art:

Ex. 1005	U.S. Patent No. 4,718,096 (“Meisel”); issued January 5, 1988; prior art under 35 U.S.C. § 102(a), (b), (e).
Ex. 1006	U.S. Patent No. 5,208,864 (“Kaneda”); issued May 4, 1993; prior art under 35 U.S.C. § 102(a), (b), (e).
Ex. 1007	U.S. Patent No. 6,148,313 (“Freidin”); filed March 30, 1998, and issued November 14, 2000; prior art under 35 U.S.C. § 102(a), (e).

Ex. 1008	U.S. Patent No. 4,239,936 (“Sakoe”), filed December 28, 1978, and issued December 16, 1980; prior art under 35 U.S.C. § 102(a), (b), (e).
Ex. 1012	Richard G. Lyons, <i>Understanding Digital Signal Processing</i> , first printed October 1996; copyrighted 1997; prior art under pre-AIA 35 U.S.C. § 102(a), (b).

Unpatentability Grounds:

Ground 1	Claims 1-2 and 4-5 are rendered obvious under 35 U.S.C. § 103 by Meisel in view of Kaneda and Freidin.
Ground 2	Claim 3 is rendered obvious under 35 U.S.C. § 103 by Meisel in view of Kaneda, Freidin, and Lyons.
Ground 3	Claim 4 is rendered obvious under 35 U.S.C. § 103 by Meisel in view of Kaneda, Freidin, and Sakoe.

II. Background

A. The '058 Patent

The '058 patent was filed May 30, 2002, claiming priority to eleven U.S. provisional patent applications, the earliest of which is U.S. Provisional No. 60/294,383, filed May 30, 2001.

The '058 patent describes a system for detecting voiced and unvoiced speech using acoustic and non-acoustic sensors. Ex. 1001, Abstract, FIGS. 1, 3. The '058 patent describes characterizing speech signals as voiced and unvoiced and, in certain embodiments, distinguishing speech signals from noise signals in speech recognition. *Id.*, 4:36-41.

The system includes a **processor (30)** connected to **acoustic microphones 10 (MIC 1 and MIC 2)** that receive acoustic signals ($m_1(n)$ and $m_2(n)$) and a **voicing sensor 20 (VAD)** that receives physiological information associated with human voicing activity, such as vocal fold vibrations. *Id.*, 2:30-37, 3:64-4:13, 5:26-29. **Voicing sensor 20** may take several forms, including an acoustic throat microphone. *Id.*, 5:26-29.

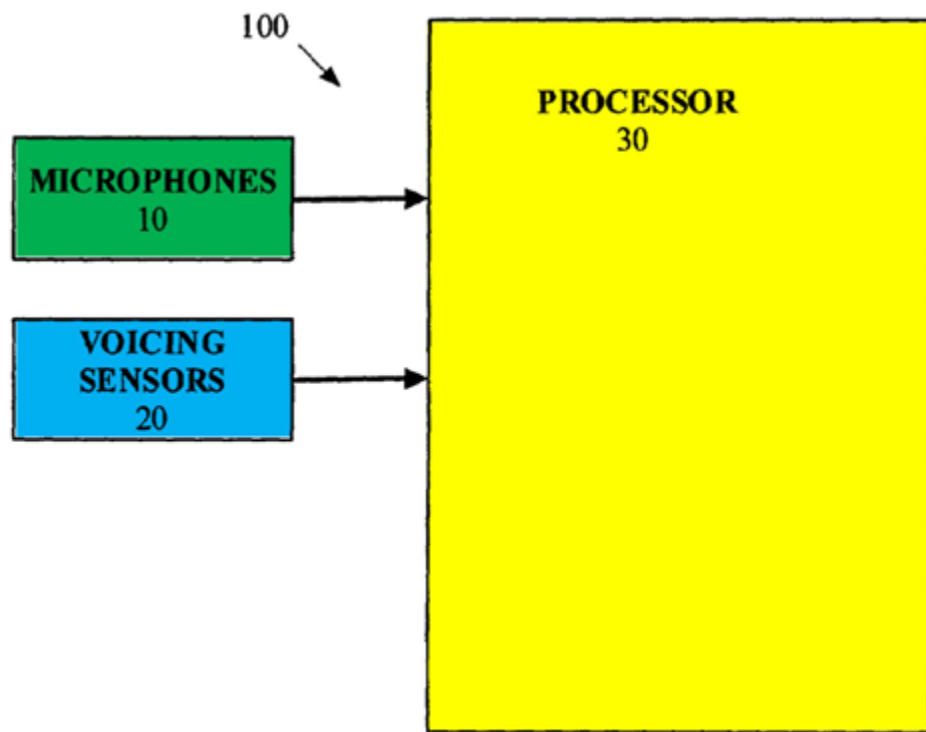


Figure 1

Id., FIG. 1 (modified, annotated).

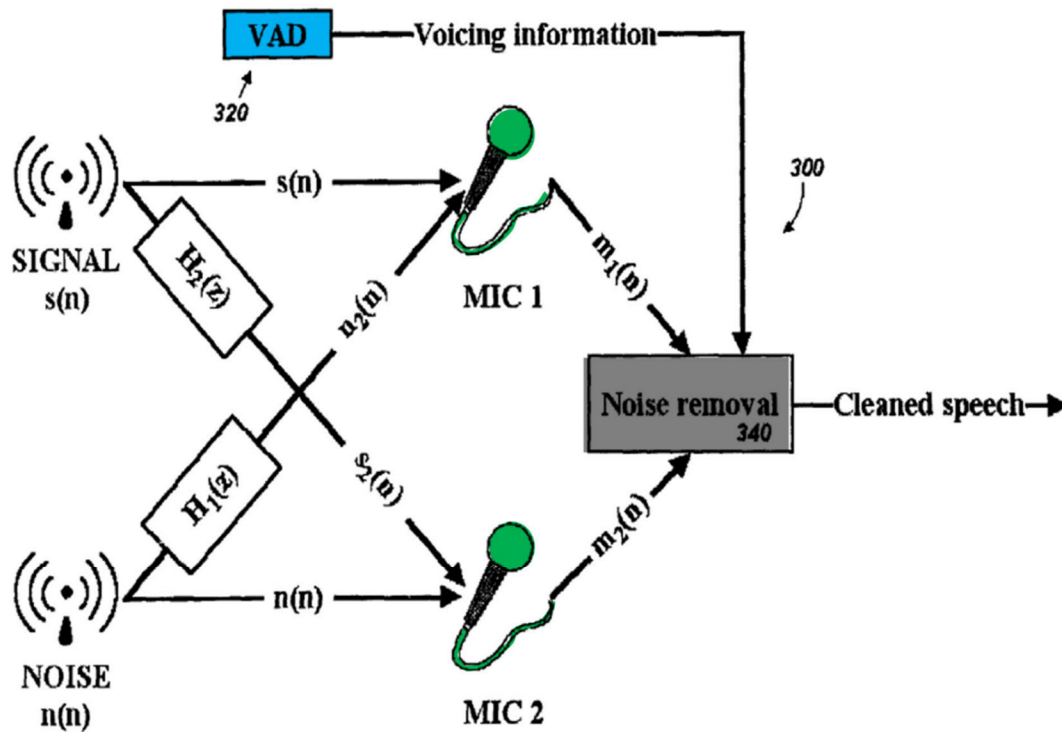


Figure 3

Id., FIG. 3 (annotated).

Processor 30 generates cross correlation data correlating the physiological information (“[v]oicing information”) from the **voicing sensor** and an acoustic signal received from one of the acoustic **microphones**. *Id.*, 4:64-5:3, 5:11-25.

Processor 30 compares the cross correlation data to an empirical correlation threshold and, if the correlation data exceeds the correlation threshold, the processor identifies information in the acoustic signal as voiced speech. *Id.*, 5:39-45. **Processor 30** also generates difference parameters, which represent a relative

difference in signal gain between portions of the acoustic signals at the two microphones, by comparing the received acoustic signals. *Id.*, 5:39-45, 7:52-67.

The processor compares the difference parameters to a predetermined gain threshold and, if the difference parameters exceed the gain threshold, the processor identifies information in the acoustic signals as unvoiced speech. *Id.*, 8:21-48. If the difference parameters are less than the gain threshold, the processor identifies that the acoustic signals as noise. *Id.*, 8:41-43. The processor may then **remove the identified noise (noise removal 340)**. *Id.*, 7:18-19.

B. Prosecution History

The '058 patent was filed as U.S. Patent Application No. 10/159,770. After the Office rejected all claims as obvious over U.S. Patent No. 5,590,241 (“Park”) and U.S. Patent No. 6,233,551 (“Cho”) (Ex. 1002, 132-134), the applicant argued that Park fails to disclose using two microphones and voicing activity detection. Ex. 1002, 124. The applicant further argued that Cho fails to disclose using two microphones and detecting voiced and unvoiced speech, as claimed. *Id.*, 125. The Office maintained the rejection of claims 1-5 as obvious over Park and Cho but withdrew the rejection of original claims 6 and 7, noting these claims were allowable. Ex. 1002, 108-110. The applicant then canceled claims 1-5 and added claims 8-10. Ex. 1002, 100. Prosecution claims 6-10 correspond to issued claims 1-

5, respectively. The '770 application issued as the '058 patent on July 17, 2007.

Ex. 1001, cover.

III. Level of Ordinary Skill in the Art

Based on the relevant factors, *In re GPAC Inc.*, 57 F.3d 1573, 1579 (Fed. Cir. 1995), a person of ordinary skill in the art (“POSITA”) would have had a minimum of a bachelor’s degree in computer engineering, computer science, electrical engineering, mechanical engineering, or a similar field, and approximately three years of industry or academic experience in a field related to acoustics, speech recognition, speech detection, or signal processing. Ex. 1003, ¶ 14-16. Work experience can substitute for formal education and additional formal education can substitute for work experience. *Id.*

IV. Claim Construction

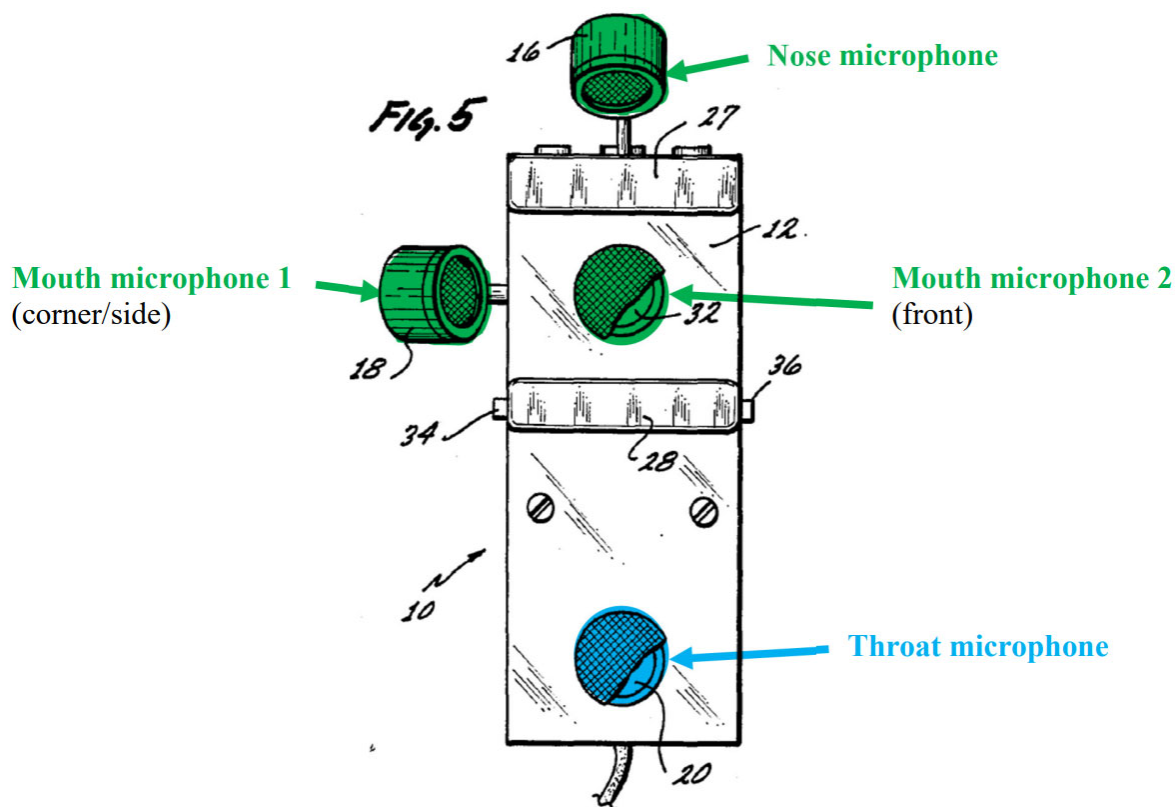
No terms need to be construed under 37 C.F.R. § 42.100(b) to resolve unpatentability. *Vivid Techs., Inc. v. Am. Sci. & Eng’g, Inc.*, 200 F.3d 795, 803 (Fed. Cir. 1999) (“[O]nly those [claim] terms need be construed that are in controversy, and only to the extent necessary to resolve the controversy”). As discussed below, the prior art is substantially the same as the '058 patent’s disclosed embodiments and, therefore, renders the claims unpatentable without claim construction.

V. The Challenged Claims Are Unpatentable and Should Be Cancelled

A. [Ground 1] Claims 1-2 and 4-5 Are Rendered Obvious by Meisel in View of Kaneda and Freidin

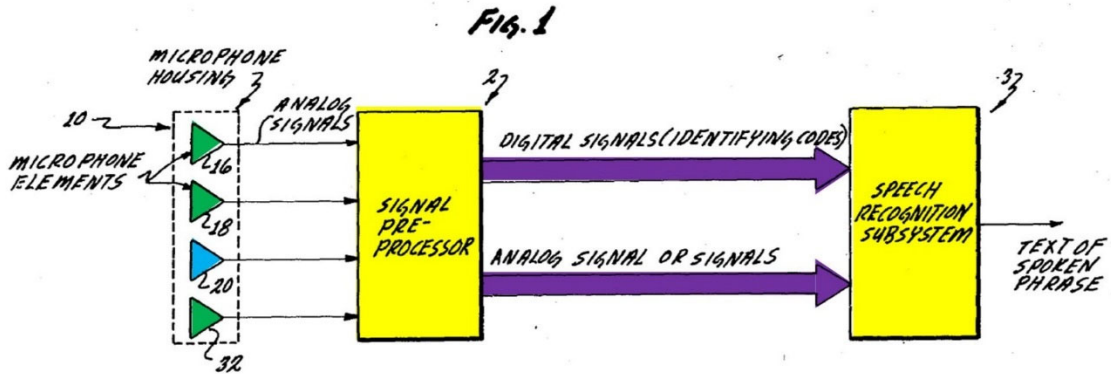
1. Meisel

Meisel discloses a speech detection system comprising a plurality of microphones and a voice activity detector. Ex. 1005, Abstract. Meisel's system includes a **nose microphone (16)** and two **mouth microphones (18, 32)** (one microphone on the corner of a speaker's mouth (18) and another microphone in front of the speaker's mouth (32)). *Id.*, 2:11-30, 4:56-59, FIG. 5. Nose microphone 16 and mouth microphones 18, 32 are acoustic microphones that detect acoustic speech signals. *Id.*, 8:49-59 (explaining that the nose and mouth microphones detect the "M" acoustic sound signals). Meisel's system includes a **throat microphone (20)** that detects physiological information associated with human voicing activity, such as "vibration of the vocal cords, or 'voicing.'" *Id.*, 2:33-35.



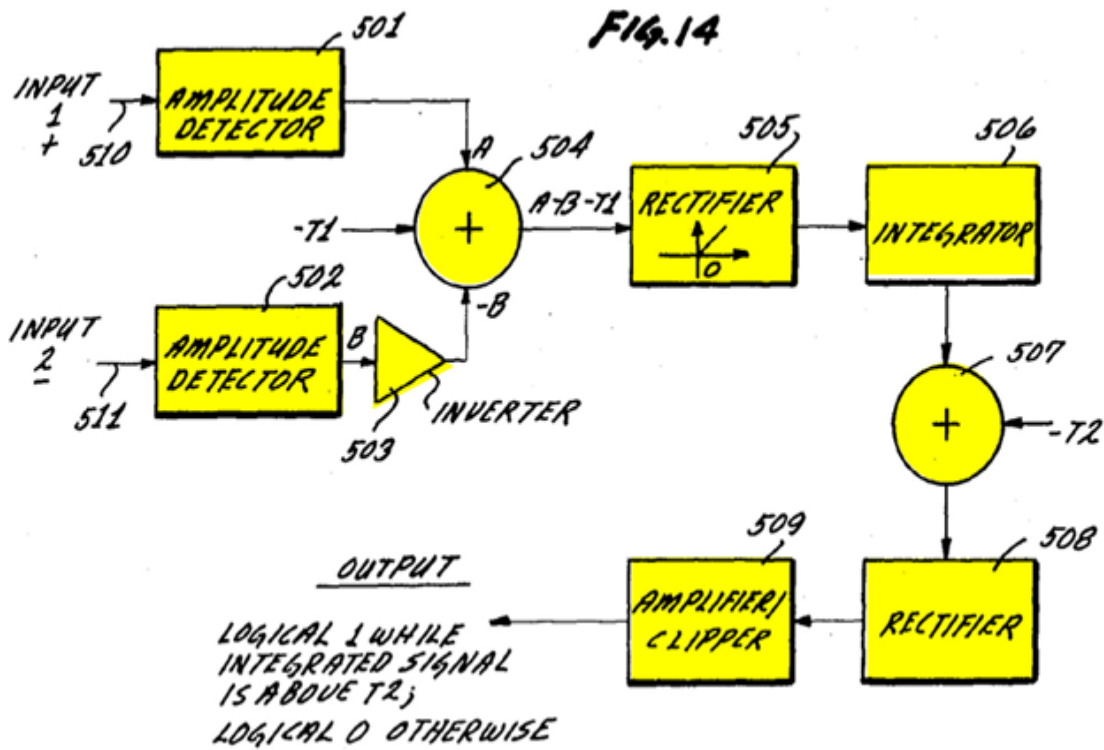
Id., FIG. 5 (annotated).

Meisel's speech recognition system includes at least one **processor**, such as **signal pre-processor (2)** and **speech recognition subsystem (3)**. *Id.*, 4:27-34. The signals from the microphones are sent to the **signal pre-processor (2)**, which compares and detects imbalances (differences) in signal amplitudes between any two microphones. *Id.*, 4:35-42. The differences are coded as Boolean variables, which are used to classify speech as voiced or unvoiced by distinguishing, for example, voiced and unvoiced stop consonants and voiced and unvoiced fricatives. *Id.*, 2:31-48.



Id., FIG. 1 (annotated).

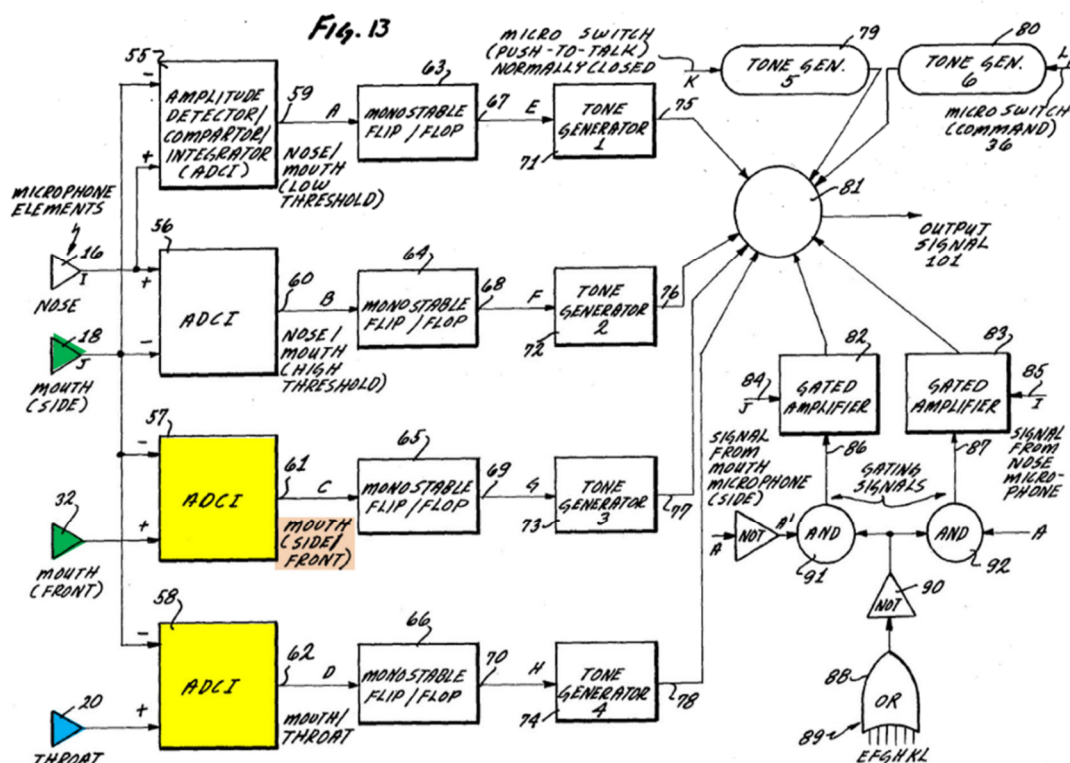
Meisel's signal pre-processor (2) uses amplitude detector/comparator/integrator units (ADCI) to distinguish between voiced and unvoiced speech using the differences in the two signals, as shown below.



Id., FIG. 14 (annotated).

For example, Meisel's signal pre-processor (2) calculates a difference between two input signals (A-B) and compares that difference to a predetermined threshold (T1). *Id.*, 9:12-29. The signal pre-processor (2) further processes this information using an integrator (506), which integrates signals over a period of time, and compares the output data to a second predetermined threshold (T2). *Id.*, 9:38-42. The signal pre-processor (2) compares the signals from the first mouth microphone (18) and the throat microphone (20) to distinguish voiced from unvoiced stop consonants and fricatives. *Id.*, 2:31-48. As one example, Meisel's speech detection system measures the difference in signal amplitudes (gains) between the received signals for the side mouth microphone (18) and the front mouth microphone (32). *Id.*, 9:38-50, 10:4-11. If the difference between the signals is greater than the threshold T1, the signal pre-processor (2) determines that the speech is voiced or unvoiced. Ex. 1003, ¶ 32; Ex. 1005, 9:38-50. Meisel provides one example of ADCI operation in Figure 14 where "the occurrence of a nasal when the input 510 is from the nasal microphone 16 and the input 511 is from the mouth microphone 18." Ex. 1005, 9:47-50. A POSITA would understand that, because Figure 14 describes the ADCI operation generally, the signal amplitude comparison step applies equally to the inputs from the side mouth microphone (18) and the front mouth microphone (32) for ADCI unit (57) in view of Figure 13. Ex. 1003, ¶ 32. Moreover, a POSITA would understand that the output of the

comparison of the side mouth microphone (18) and the front mouth microphone (32) in ADCI unit (57) would be combined with the output of the throat microphone (20) to distinguish voiced and unvoiced speech, further refining the speech recognition process. Ex. 1005, 2:31-48 (explaining that the throat microphone (20) would be used in combination with other microphones to distinguish voiced and unvoiced speech); Ex. 1003, ¶ 32. Meisel explains that “[a] similar argument applies to distinguishing voiced fricatives (v, z, and th) from unvoiced fricatives (f, s, and sh).” Ex. 1005, 2:46-48. Thus, a POSITA would recognize that Meisel’s voicing detection (throat microphone (20)) and fricative detection (ADCI unit (57)) would be combined to specify whether the fricative is voiced or unvoiced. Ex. 1003, ¶ 32.

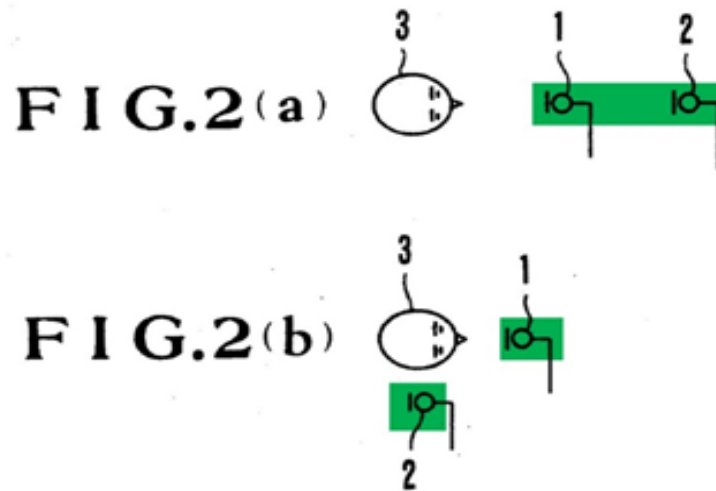


Ex. 1005, FIG. 13 (annotated).

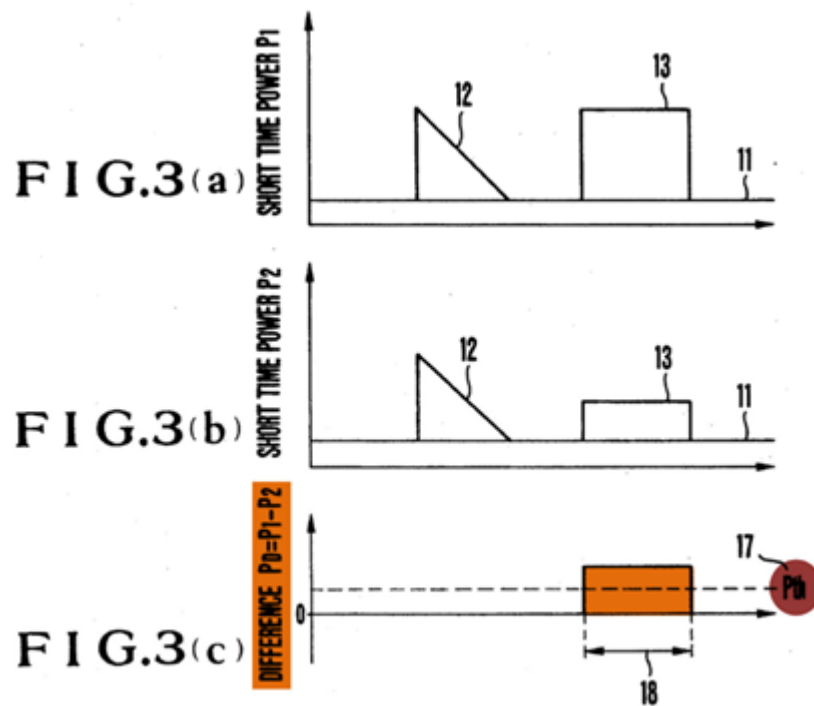
Meisel further discloses that its speech detection system operates in conditions with varying levels of background noise, such as “under conditions where there is background noise and where the speaker can continue to perform other duties,” suggesting that it operates in environments where noise is present and should be removed in processing. Ex. 1005, 1:10-15, 1:46-50. In Figure 14, if signal A does not exceed signal B by a threshold amount, T1 (“A-B-T1” in Figure 14 is greater than zero), the summing unit 504 does not provide an output, rejecting the signals as noise. Ex. 1003, ¶ 33; Ex. 1005, 9:12-25.

2. Kaneda

Kaneda discloses a conventional speech recognition system including **two acoustic microphones**. Ex. 1006, 1:55-57, FIG. 2(a). Speech signals received at microphone 1 result in a higher power output than those received at microphone 2. *Id.*, 1:61-63. This is shown in Figures 3(a) and 3(b), where the speech signal power (13) of microphone 1 (P_1) is greater than that of microphone 2 (P_2). *Id.*, 2:1-12.



Id., FIGS. 2(a)-2(b) (annotated).



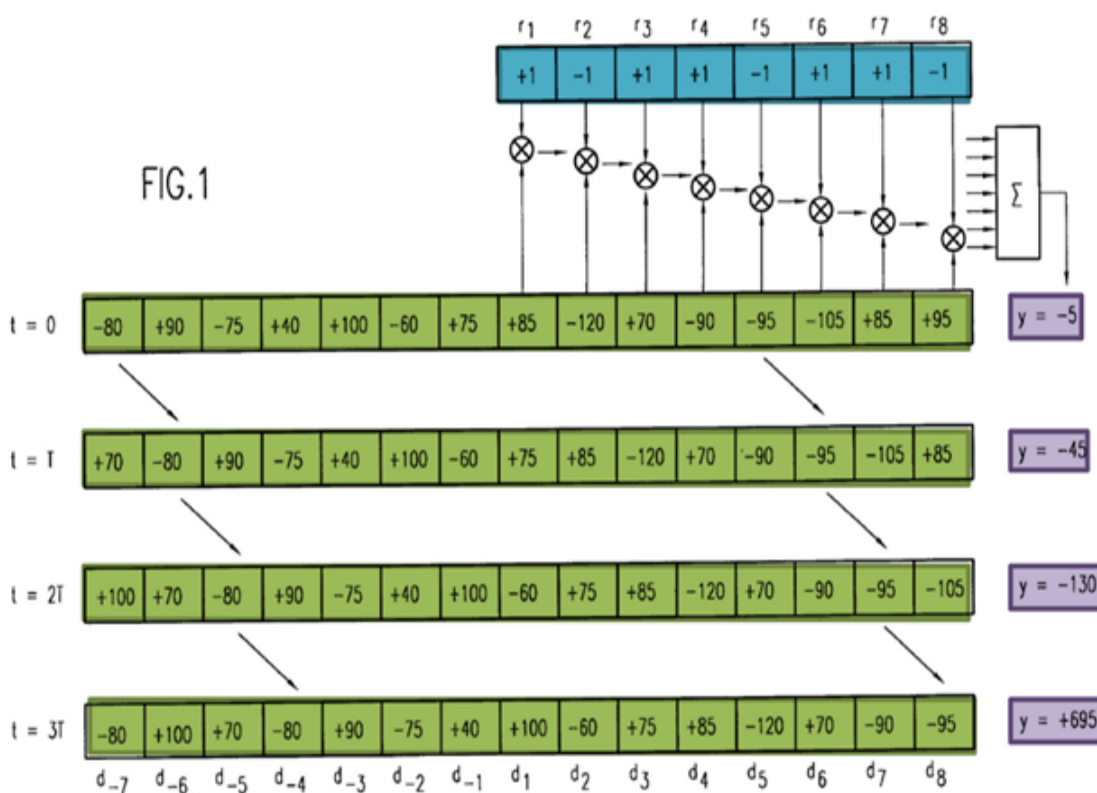
Id., FIGS. 3(a)-3(c) (annotated).

This “conventional” speech recognition system calculates a **power difference PD** between the power outputs of the two microphones, “PD (=P1-P2).” *Id.*, 2:12-15. “When the **power difference PD** is larger than a given **threshold value Pth17**, a corresponding time period 18 is detected as a speech period.” Ex. 1006, 2:12-18. Similarly, when PD is less than Pth17, the corresponding time period is noise. Ex. 1003, ¶ 35; Ex. 1006, 2:18-22.

3. Freidin

Freidin discloses a correlator that determines a correlation between digital signals in “a number of applications including speech recognition[.]” *Id.*, 1:7-26.

Freidin Figure 1 shows, for example, “a communication receiver correlator operating with a **correlation reference sequence**, which is eight (8) bits in length, to produce a correlation peak at the time when the **received data sequence** matches the **reference sequence**.” *Id.*, 1:37-41.



Id., FIG. 1 (annotated).

Freidin discloses that “[a]t each clock period, the last eight samples of input data d_i ($i=1,8$) are multiplied sample by sample against the reference sequence bits r_i ($i=1,8$) to produce a **correlation value y** .” *Id.*, 1:43-46. Freidin also discloses that the “correlation value y is compared against a **correlation threshold, y_{TH}** , and

whenever y exceeds y_{TH} , a correlation match is declared indicating that the correlation sequence has been received and the receiver is thereby synchronized to the transmitter.” *Id.*, 1:48-52. Freidin also explains that “many variations are possible which remain within the concept and scope of the invention,” not to the exact embodiments disclosed. *Id.*, 11:49-12:4.

4. Independent Claim 1

Meisel teaches or suggests each element of claim 1. Kaneda confirms and renders obvious calculating difference parameters between the acoustic signals received at each of the two receivers, as recited in [1a]. Freidin confirms and renders obvious generating cross correlation data between physiological information and acoustic signals, as recited in [1d], and the cross correlation data exceeding a correlation threshold, as recited in [1e].

a. **[1Preamble] “A system for detecting voiced and unvoiced speech in acoustic signals having varying levels of background noise, comprising:”**

To the extent the preamble is limiting,¹ Meisel discloses [1Preamble]. Ex. 1003, ¶ 39. Meisel discloses using a throat microphone, a nose microphone, and one or more mouth microphones to detect and classify a speaker’s speech as

¹ Petitioner takes no position on whether preambles are limiting. If limiting, the preambles are taught or rendered obvious.

voiced or unvoiced when the signals are processed by the pre-processing system.

Ex. 1005, 2:11-48, 2:61-3:4, 10:6-11, 13:10-22, 14:57-61.

Meisel's system compares "the signal from the throat microphone [20] to that of the mouth microphone [18]" to "distinguish, for example, voiced from unvoiced stop consonants" and "voiced fricatives (v, z, and th) from unvoiced fricatives (f, s, and sh)." Ex. 1005, 2:31-53. A POSITA would understand that distinguishing voiced from unvoiced stop consonants and voiced from unvoiced fricatives are examples of distinguishing voiced from unvoiced speech. Ex. 1003, ¶ 40, Ex. 1005, 2:35-45 ("By comparing the signal from the throat microphone to that of the mouth microphone, one can distinguish, for example, voiced from unvoiced stop consonants," e.g. "unvoiced stops (p, t, and k)" where "the vocal cords do not vibrate" from "voiced stops (b, d, and g)" where "the vocal cords vibrate during the closure.").

Meisel further discloses that the sounds received from the nose and mouth microphones are acoustic signals. Ex. 1003, ¶ 41; Ex. 1005, Claim 6 (microphones "for detecting acoustic sounds substantially emanating from a corner of a speaker's mouth" and "acoustic sounds substantially emanating from the speaker's nose"), 10:22-29.

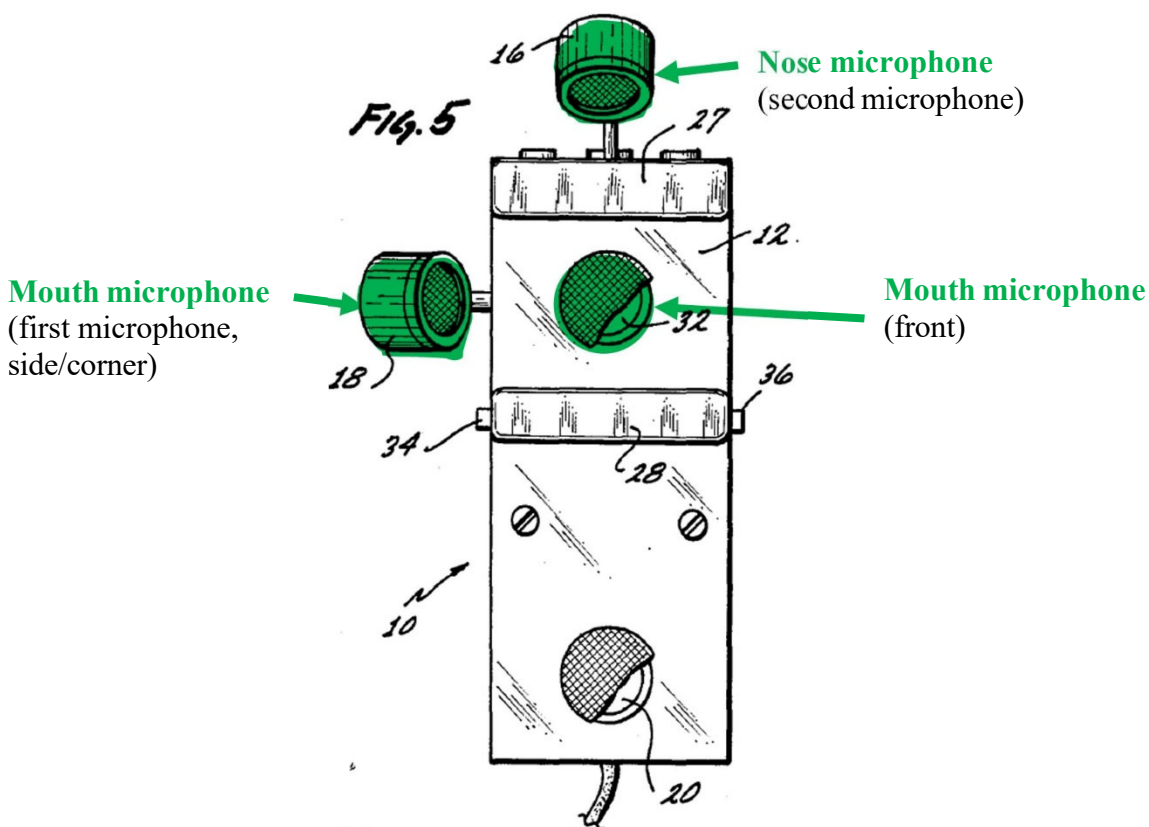
Meisel suggests that the acoustic signals have varying levels of background noise. Ex. 1003, ¶ 42. Meisel explains that "close-talking" microphones, which

receive speech signals as well as noise (e.g., background noise), were known in the prior art, suggesting that conventional microphones known at the time received different levels of noise along with speech. *Id.*; Ex. 1005, 2:54-60. Meisel identifies several problems with “background noise” making it difficult for prior art systems to distinguish speech. Ex. 1005, 1:7-50. A POSITA would, therefore, understand that Meisel’s system operates in environments having varying background noise. Ex. 1003, ¶ 42.

Moreover, “varying levels of background noise,” would have been obvious in view of Kaneda. Ex. 1003, ¶ 43. Kaneda, like Meisel, is directed to a speech detection system including two microphones. *Supra* §§ V.A.1-V.A.2. A POSITA would have looked to Kaneda to understand the operation of the acoustic microphones in Meisel’s system (e.g., nose and mouth microphones). Ex. 1003, ¶ 44; Ex. 1005, 1:7-16, 2:11-53. Kaneda discloses that microphone signals “generally contain[] stationary noise 11 (noise having almost a constant power, such as air-conditioning noise or fan noise of equipment), unstationary noise 12 (noise whose power is greatly changed, such as a door closing sound and undesired speech), and desired speech 13.” Ex. 1006, 1:30-35. A POSITA would understand that stationary and unstationary noise refers to varying levels of background noise that are commonplace in environments for Meisel’s system, such as offices. Ex. 1003, ¶ 44.

b. [1a] “at least two microphones that receive the acoustic signals;”

Meisel discloses three microphones that receive the acoustic signals. Meisel discloses a **nose microphone (16)** that detects acoustic sounds emanating from the speaker’s nose and two **mouth microphones (18, 32)** that detect acoustic sounds emanating from the speaker’s mouth. Ex. 1005, FIG. 5, Abstract, 5:29-36. Meisel explains that a **first microphone (mouth microphone 18)** “is disposed so as to detect sounds substantially emanating from a corner of the speaker’s mouth” and a **second microphone (nose microphone 16)** “is disposed so as to detect sounds substantially emanating from the nose of the speaker.” Ex. 1005, 2:11-17.



Ex. 1005, FIG. 5 (annotated).

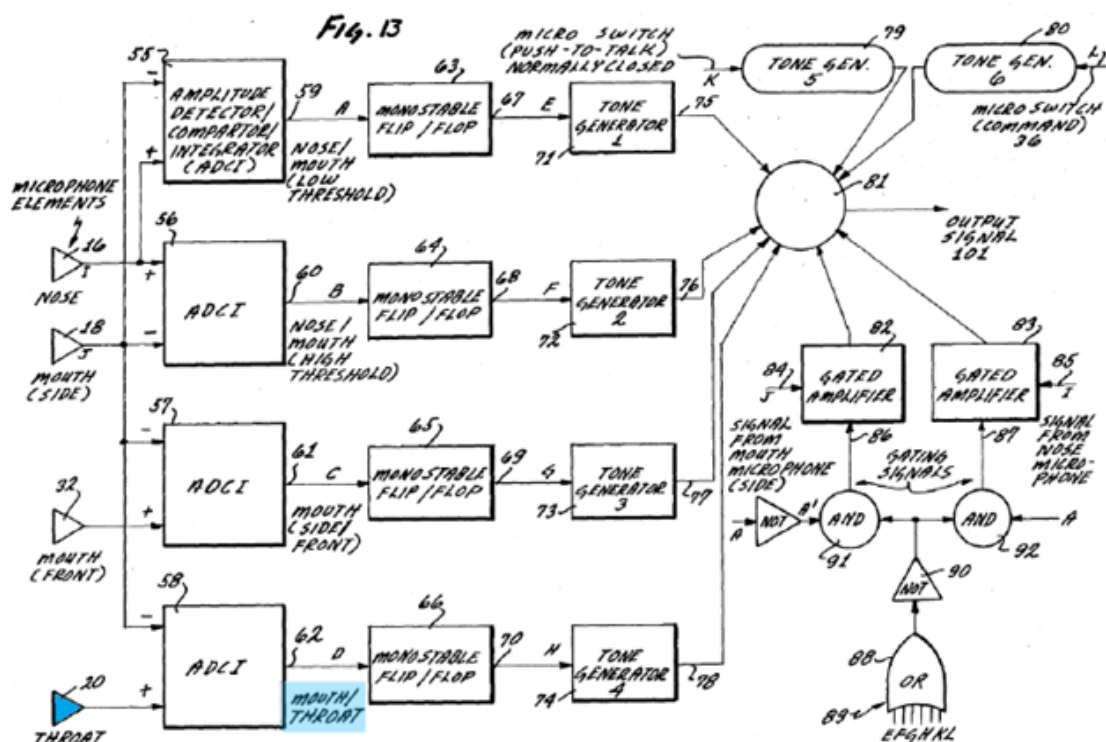
The nose microphone (16) and the mouth microphones (18, 32) are acoustic microphones that receive acoustic signals from the speaker's nose and mouth, respectively. Ex. 1003, ¶ 46. Meisel discloses that the mouth microphones 18, 32 are a “first means for detecting *acoustic sounds* substantially emanating from the mouth of the speaker and for producing signals representing such sounds,” and that the nose microphone 16 is the “second means for detecting *acoustic sounds* substantially emanating from the nose of the speaker and for producing signals representing such sounds.” Ex. 1005, Claim 1 (emphases added), 5:29-36; Ex. 1003, ¶ 46. Meisel explains that the nose microphone 16 detects the nasal acoustic signals or “the nasal component in order to provide a reliable recognition of speech content.” Ex. 1005, 8:49-59. Meisel further explains that the throat microphone (20) and the second mouth microphone (32) detect “the sounds emanating from the throat or immediately in front of the mouth for certain words in order to provide a reliable recognition of such spoken words.” Ex. 1005, 8:55-59.

c. [1b] “at least one voicing sensor that receives physiological information associated with human voicing activity; and”

Meisel discloses a **voicing sensor** that receives physiological information associated with human voicing activity, such as vibrations from the user's throat via acoustic throat microphone 20.

The '058 patent's voicing sensors include "any sensor that is able to detect human tissue motion associated with speech" (Ex. 1001, 3:61-63), including an "acoustic throat microphone[]," such as Meisel's throat microphone. Ex. 1001, Claim 5. Meisel states that its speech recognition system is "usable in the normal work environment of the speaker, particularly under conditions where there is background noise and where the speaker can continue to perform other duties." Ex. 1005, 1:10-15.

Meisel discloses an **acoustic throat microphone (20)**, the same as the '058 patent, which detects vocal cord vibrations associated with voicing activity (e.g., voiced speech) through sounds emanating from the speaker's throat (physiological information). Ex. 1005, 2:31-35; Ex. 1003, ¶ 47. In particular, Meisel explains that "[t]he throat microphone, if used, detects the vibration of the vocal cords, or 'voicing,'" and "the [throat] microphone 20 is disposed to detect vibration of the vocal cords through sounds emanating from the throat of a speaker." Ex. 1005, 2:31-35, 5:29-36. A POSITA would have recognized that detecting "vibration of the vocal cords" is receiving physiological information about "voicing" activity of the speaker. Ex. 1003, ¶ 49; Ex. 1005, 2:33-35.



Ex. 1005, FIG. 13 (annotated).

- d. [1c] “at least one processor coupled among the microphones and the voicing sensor,”

Meisel discloses at least one processor coupled among the microphones and the voicing sensor. Ex. 1003, ¶ 50. Meisel discloses the same system architecture as the '058 patent for coupling a processor among the microphones and the voicing sensor. Figure 1 of the '058 patent shows **microphones (10)** and **voicing sensors (20)** connected to the same **processor (30)**. Ex. 1001, FIG. 1.

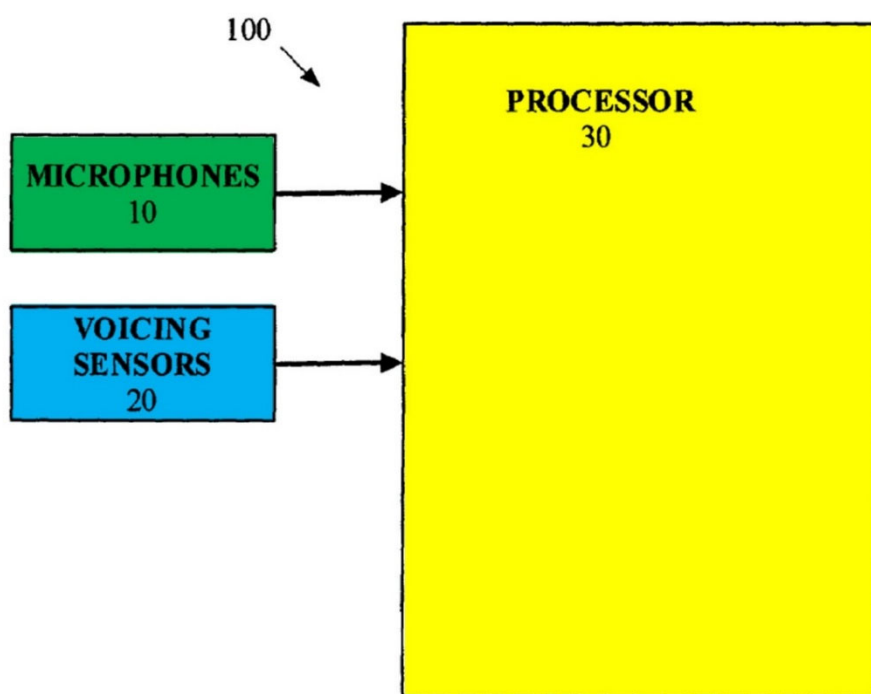
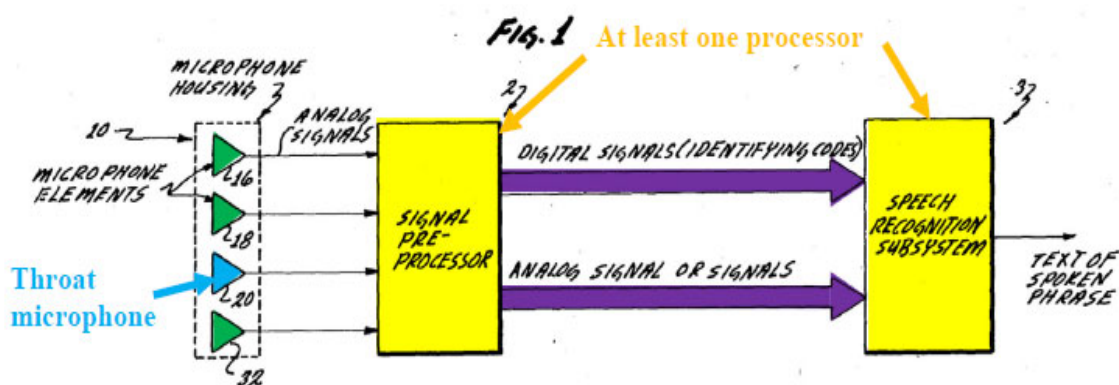


Figure 1

Ex. 1001, FIG. 1 (modified, annotated).

Meisel's Figure 1 similarly shows that all microphones, including the **nose microphone (16)** and **two mouth microphones (18, 32)**, and the **voicing sensor (throat microphone (20))** are all connected, or coupled, to the same **signal pre-processor (2)** and then to the **speech recognition subsystem (3)**. Ex. 1005, FIG. 1, 4:27-34.



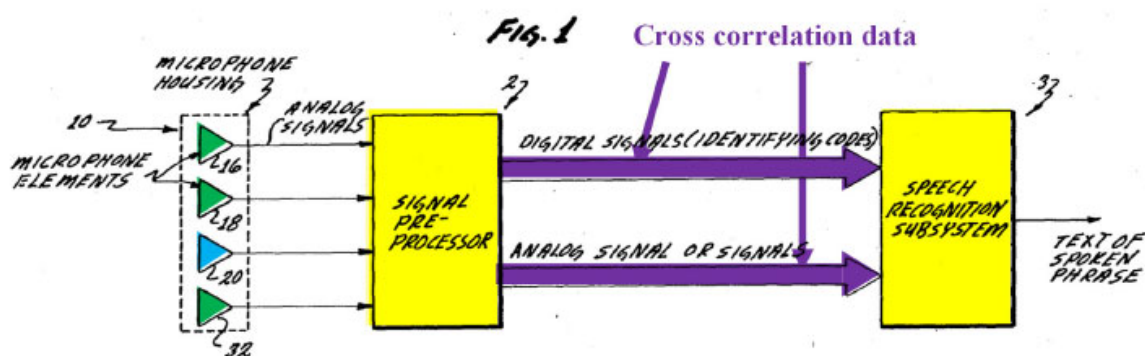
Ex. 1005, FIG. 1 (annotated).

Meisel discloses that the signals from the **acoustic microphones (nose microphone (16) and mouth microphones (18, 32))** and the **voicing sensor (20)** are sent to the **signal pre-processor (2)** to detect differences in the signal amplitudes. Ex. 1005, 4:35-42 (“The pre-processor 2 processes the signals from the microphones 16, 18, 20 and 32 to determine at each instant imbalances in the relative amplitude of each microphone. These imbalances are indicative of the production of certain speech sounds.”). The differences in the signal amplitudes are coded as Boolean variables and then sent to the speech recognition subsystem (3), which processes the codes to detect and classify speech. Ex. 1005, 4:39-53 (“The detected imbalances are coded as Boolean variables in a conventional manner to create a digital code corresponding to the pattern of imbalances.”), 3:33-41 (“The selected signals and the codes are passed to the recognition subsystem. The

selected signals and the codes may be processed to detect and classify syllables of speech, and the identified syllables may be combined into phrases or sentences.”).

- e. [1d] “wherein the at least one processor; generates cross correlation data between the physiological information and an acoustic signal received at one of the two microphones;”

Meisel discloses generating cross correlation data between the physiological information and an acoustic signal received at one of the two microphones. Ex. 1003, ¶ 53. The **nose and mouth microphones (16, 18, 32)** receive acoustic signals, as explained in [1a] above (*supra* § V.A.4.b), and the **throat microphone (20)** receives physiological information, as explained in [1b] above (*supra* § V.A.4.c). Meisel further discloses that the processor (e.g., **signal pre-processor (2)**) generates **cross correlation data** between the physiological information and an acoustic signal received at one of the two microphones.



Ex. 1005, FIG. 1 (annotated).

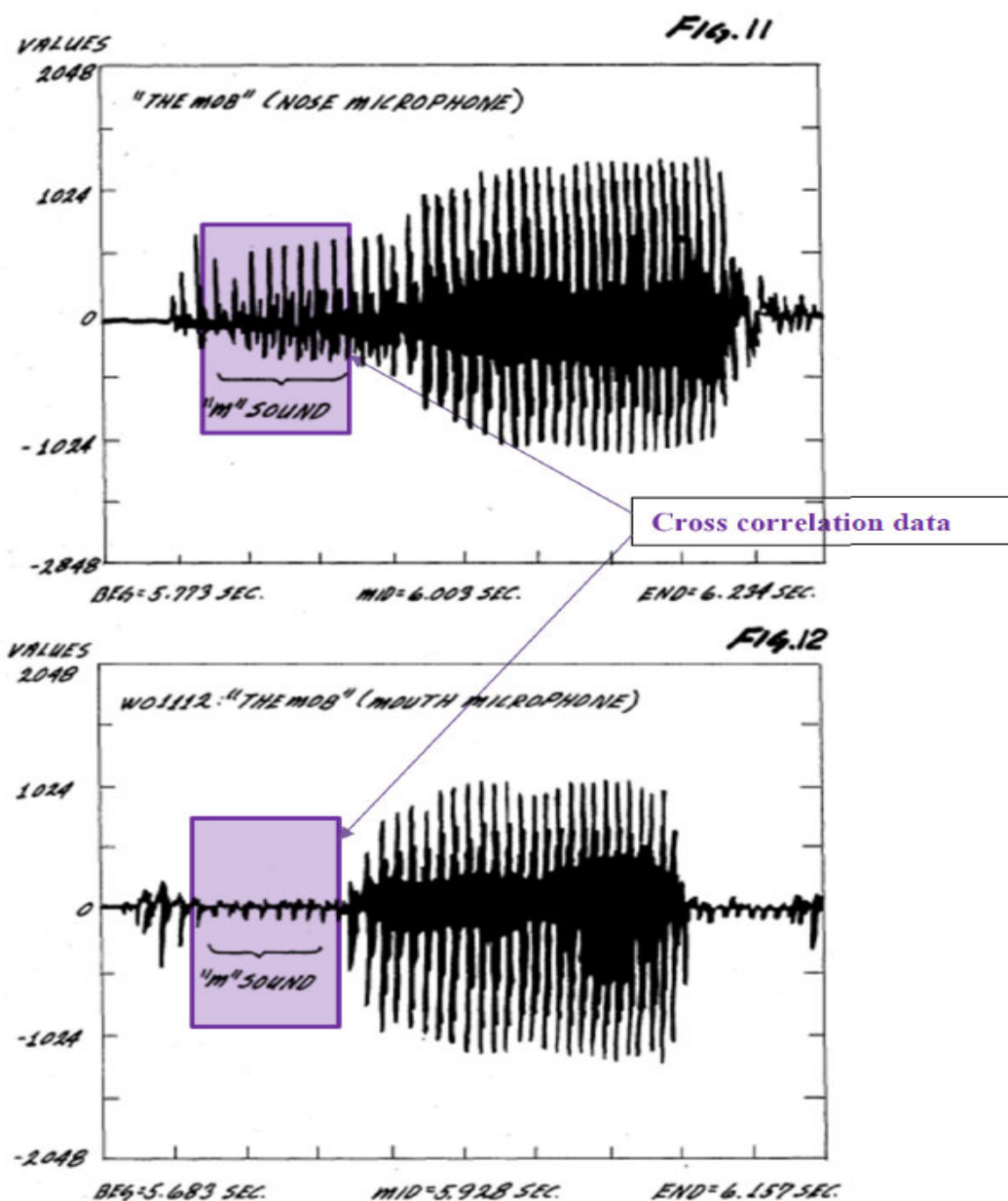
Meisel discloses correlating the acoustic information from at least one of the mouth microphones (e.g., **the first mouth microphone (18)** located to the side of the speaker's mouth) with the physiological information from the **throat microphone (32)** by comparing their signals to detect differences in their signal amplitudes. Ex. 1005, 4:35-42; Ex. 1003, ¶ 54.

As explained in [1c] (*supra* § V.A.4.d), Meisel discloses sending signals from the **nose and mouth microphones (16, 18, 32)** and the **throat microphones (20)** to the **signal pre-processor (2)**, which compares these signals and determines “imbalances” (differences) in the amplitude of signals from different microphones. Ex. 1005, 4:35-42, FIG. 1. These differences in signal amplitudes are then coded as “the pattern of imbalances” and sent to the **speech recognition subsystem (3)**. Ex. 1005, 4:35-44, FIG. 1. The speech recognition subsystem (3) uses the coded differences in signal amplitudes to complete speech recognition. Ex. 1005, 4:43-53. A POSITA would understand that a pattern of imbalances is a type of **cross correlation data** because it contains, or is based on, a correlation between signal amplitudes of the acoustic microphones (e.g., **side mouth microphone 18**) and the **throat microphone 20**. Ex. 1003, ¶ 55.

Meisel's speech recognition system compares, for example, “the signal from the throat microphone to that of the mouth microphone” to “distinguish, for example, voiced from unvoiced stop consonants” and “voiced fricatives (v, z, and

th) from unvoiced fricatives (f, s, and sh).” Ex. 1005, 2:31-48. As explained in [1a] (*supra* § V.A.4.b), the nose microphone (16) and the mouth microphones (18, 32) receive acoustic signals emanating from the speaker’s nose and mouth. Ex. 1005, 2:11-17.

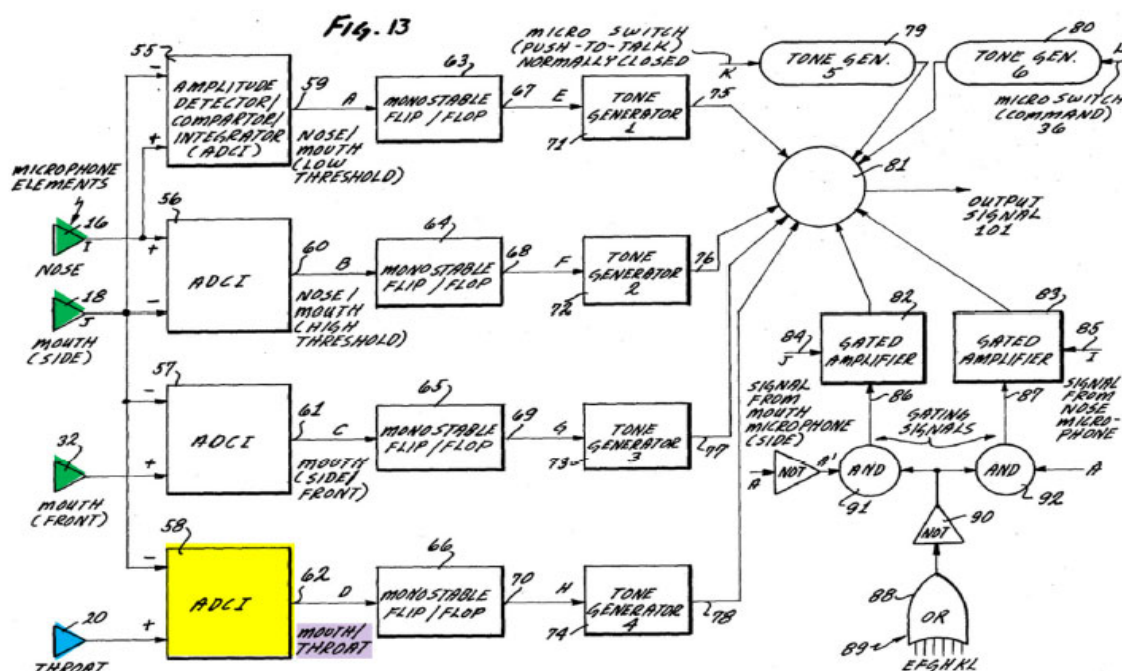
In Meisel Figures 11 and 12, the signals from the nose microphone (16) and the mouth microphones (18, 32) are correlated to determine whether a portion of the signals is nasal. Ex. 1005, 8:24-59, FIGS. 11-12. While Figures 11 and 12 illustrate cross correlation of the signals from the nose microphone (16) and the first mouth microphone (18), Meisel explains that “[s]imilar proof can be provided of the importance of recording the sounds emanating from the throat or immediately in front of the mouth for certain words in order to provide a reliable recognition of such spoken words.” Ex. 1005, 8:55-59. Indeed, Meisel’s Figure 13 shows an ADCI for comparing different signals, such as ADCI unit (58) for comparing the throat microphone (20) with the side mouth microphone (18). Ex. 1005, FIG. 13, 9:60-66. Therefore, a POSITA would understand that the cross correlation process in Figures 11 and 12 applies equally to the signals from the throat microphone and, for example, the side mouth microphone. Ex. 1003, ¶ 57.



Ex. 1005, FIGS. 11-12 (annotated).

In Figure 13, the first mouth microphone (18) and the throat microphone (20) are connected to ADCI unit (58), which is part of the signal pre-processor (2). Ex. 1005, 9:60-66. The output (62) from the ADCI unit (58) "will become logically true when a voiced stopped sound is being pronounced." *Id.*, 10:9-11. In

other words, if the output of the ADCI unit (58) (the comparison of a throat signal and side mouth signal) is non-zero, then the signal pre-processor (2) determines that a voiced speech is present. Ex. 1005, 9:12-22; Ex. 1003, ¶ 58. A POSITA would understand that the throat microphone (20) generally indicates whether the corresponding speech types (e.g., fricative, plosive, stopped) produced at any of the other microphones (16, 18, 32) are voiced or unvoiced. Ex. 1003, ¶ 58.



Ex. 1005, FIG. 13 (annotated).

Moreover, generating the claimed cross correlation data would have been obvious to a POSITA by cross-correlating Meisel's throat microphone (20) and sound received at one or more of the acoustic microphones (16, 18, 32) to identify

voiced speech because cross correlation was a well-known and routine process in the art, as taught by Freidin. A POSITA would have found this cross correlation obvious to identify sounds at the throat and mouth microphones as their respective voiced or unvoiced type based on whether the throat microphone has a high signal amplitude, rather than rejecting them as noise. Such signal comparison would have been understood by a POSITA to allow for more accurate characterization of Meisel's speech. Ex. 1003, ¶ 59.

Freidin discloses a correlator that correlates an input signal sequence, d_i , with a reference signal sequence, r_i , to calculate a correlation value, y , using the formula $y=d_i*r_i$ for the sequence $i=(1,n)$. Ex. 1007, 1:10-26, 1:37-52. A POSITA would have recognized that Freidin's correlation formula is a cross correlation formula. Ex. 1003, ¶ 60; *see, e.g.*, Ex. 1016, 20 ("Given two real-valued sequences $x(n)$ and $y(n)$ of finite energy, the *crosscorrelation* of $x(n)$ and $y(n)$ is a sequence $r_{xy}(l)$ defined as $r_{x,y}(l) = \sum_{n=-\infty}^{\infty} x(n) y(n - l)$."). A POSITA would have found it obvious to apply Freidin's correlator to the input signals from Meisel's side mouth microphone (18) and throat microphone (20) to calculate a correlation value (cross correlation data) between the two input signals to identify voiced speech when the signals are both present at both microphones (e.g., the difference would be small because a signal is present both at the throat and mouth microphones). Ex. 1003, ¶ 60.

i. Motivation to combine

A POSITA would have been motivated to combine Meisel's speech processing system with Freidin's correlator to calculate cross correlation data of the signals received at Meisel's microphones. Ex. 1003, ¶ 61. Like Meisel (*infra* § V.A.4.g), Freidin discloses comparing two received signals (a reference signal sequence and an input signal sequence) to calculate cross correlation data and determine whether the signals are correlated. Ex. 1007, 1:43-46. Freidin describes the cross correlation function to produce the cross correlation data, y, to identify a particular result, such as whether Meisel's speech is voiced speech. Ex. 1007, 1:48-52; Ex. 1003, ¶ 61. A POSITA would have looked to Freidin to implement the cross correlation and speech identification performed by Meisel's ADCI units and signal pre-processor (2) and speech recognition subsystem (3) because Meisel describes a system in which the signals are compared and correlated with each other and produce a result, but Freidin provides a well-known cross correlation framework to render Meisel's identification more objective and accurate. Ex. 1003, ¶ 62. A POSITA also would have been motivated to use cross correlation, as disclosed in Freidin, to identify voiced speech types because such voiced speech types correlate when the signals at the throat microphone and signal microphone (e.g., mouth microphone) are high and the differences of the ADCI calculation may therefore be small. Ex. 1003, ¶ 62. Thus, a POSITA would have been motivated to

use Freidin's cross correlation calculation to improve the accuracy of voiced and unvoiced speech identification already performed by Meisel. *Id.*; *see, e.g.*, Ex. 1005, 2:31-48.

A POSITA would also have been motivated to use Freidin's cross correlation to identify voiced speech in Meisel because Freidin discloses that "[c]orrelators have been widely used in signal processing to measure the degree of correspondence between a data sequence and a predetermined pattern or reference sequence," including "in a number of applications including speech recognition." Ex. 1007, 1:7-26; Ex. 1003, ¶ 63. A POSITA would have recognized that Meisel's signal processing techniques would be most easily implemented using digital processors, and Freidin provides the guidance needed to find signal correlations.

ii. Expectation of success

A POSITA would have reasonably expected to succeed in combining Meisel's speech recognition system and Freidin's correlator for correlating received speech signals because the combination uses only known techniques of standard signal processing functions to achieve predictable results. Ex. 1003, 64. A POSITA would have recognized, and it would have been well within a POSITA's skill, to implement Meisel's processing in digital processors; in fact, it would have been most straightforward to process the received signals using digital processors. Ex. 1003, ¶ 64. It was well-known at the time to cross-correlate received signals to

determine signal correlation. *Id.*; Ex. 1013 (Rainton), Abstract (“a cross correlator calculates a cross correlation value by using a predetermined cross correlation function based on the filtered first and second signals”); Ex. 1015 (Matsuo), 7:63-8:26 (discussing using a cross correlation and cross correlation coefficient (R) to determine the correlation between two signals); Ex. 1014 (Chu), Abstract, 2:6-19; Ex. 1016 (Ingle), 20, 29; Ex. 1017 (Byers), 7:40-48, 12:33-56; Ex. 1018 (Makino), Abstract, 7:2-9; Ex. 1015 (Matsuo), 3:10-15; Ex. 1020 (Matsuo 2), 17:1-16; Ex. 1021 (Benesty), Abstract, 4:4-13, 6:16-9:35. Indeed, cross correlation calculations on two received signals were so conventional that it was a built-in function for signal processing programs, such as the XCORR function of Matlab.² *See, e.g.*, Ex. 1009, 6-367 through 6-372.

Indeed, cross correlation calculations date back decades to at least 1966 with a book stating that “[t]he optimal detection scheme for the signal specified exactly is therefore cross correlation between the expected signal and the received waveform.” Ex. 1024 (Green), 163; *see also* Ex. 1023 (Duda), 279-80 (1973 reference explaining “template matching” using cross correlation to determine

² Cross correlation functions were so conventional that the ’058 patent identifies the built-in conventional XCORR function rather than explaining cross correlation. Ex. 1001, 4:64-5:3; Ex. 1003, ¶ 64.

similarity in a template and a region of interest in an image.). In fact, cross correlation calculations were standard calculations for undergraduate signal processing classes at the earliest claimed priority date of the '058 patent. Ex. 1003, ¶ 65. Thus, applying Freidin's cross correlation function to compare Meisel's acoustic and throat microphone signals (or any other received microphone signals) to correlate the signals and identify, for example, voiced speech, was nothing more than the application of conventional, decades-old processing techniques applied in their ordinary and expected manner to achieve the expected result. Ex. 1003, ¶ 65. A POSITA would have found that implementing Freidin's cross correlation in Meisel's speech recognition system was combining prior art elements according to known methods (digital signal processing of speech) to yield predictable results (improving the accuracy of speech detection). Ex. 1003, ¶ 65. Further, combining Meisel and Freidin would have been applying a known technique (a correlator) to a known device (a speech recognition system) ready for improvement to yield predictable results (improving the accuracy of speech detection).

The '058 patent discloses using "cross-correlation (XCORR) between acoustic signal from microphone 1 and the sensor data," and that the thresholds are empirically determined. Ex. 1001, 4:66-5:3, 5:42-44. "XCORR" was a well-known signal abbreviation and is a common function name for cross correlation functions in signal processing software, such as MATLAB, at least as early as 1999. Ex.

1003, ¶ 66; Ex. 1009, 3-2 through 3-3. For example, the MATLAB Manual for The MathWorks' Signal Processing Toolbox explains that "[t]he functions xcorr and xcov estimate the cross-correlation and cross-variance sequences for random processes." Ex. 1009, 3-2, 3-4. It also provides that "the true cross-correlation sequence is a statistical quantity defined as $\gamma_{xy}(m) = E\{x_n y^{*}_{n+m}\}$ where x_n and y_n are stationary random processes, $-\infty < n < \infty$, and $E\{\}$ is the expected value operator." Ex. 1009 at 3-2, 3-4. A POSITA would understand that speech signals are modeled as a series of stationary random processes over a period such as about 20ms. Ex. 1025 (Deller), 231; Ex. 1003, ¶ 66. A POSITA would have known how to apply cross correlation equations because they are conventional functions, such as one built into MATLAB, using the signals from the acoustic microphones (e.g., first mouth microphone (18) and the throat microphone (20)) as the correlation inputs to arrive at the cross correlation between the two microphones. Ex. 1003, ¶ 66. For example, as explained above, Figure 13 in Meisel shows ADCI unit (58), which compares the signals received at the throat microphone (20) with the signals received at the side mouth microphone (18) for correlation. Ex. 1005, FIG. 13, 9:60-66. Based on this disclosure, a POSITA would have been motivated to look at the MATLAB Manual because the Signal Processing Toolbox was a commonly used program in the art for speech signal processing at the time. Ex. 1003, ¶ 66; Ex. 1016, 29. Given that both Meisel and Kaneda use signal processing in their

speech recognition systems, a POSITA would have had a reasonable expectation of success in implementing the cross correlation (e.g., xcorr) function as a conventional signal processing operation to determine cross correlation data between Meisel's throat microphone signal and one or more of the mouth or nose microphone signals. Ex. 1003, ¶ 66.

- f. [1e] **“identifies information of the acoustic signals as voiced speech when the cross correlation data corresponding to a portion of the acoustic signal received at the one receiver exceeds a correlation threshold;”**

Meisel discloses that its processor (**signal pre-processor (2)** and **speech recognition subsystem (3)**) identifies information of the acoustic signals as **voiced speech** when the cross correlation data corresponding to a portion of the acoustic signal received at the one receiver exceeds a correlation threshold. Ex. 1003, ¶ 67. As explained in [1d] (*supra* § V.A.4.e), Meisel discloses **a signal pre-processor (2)** that receives acoustic signals from the **nose microphone (16)** and the **mouth microphones (side (18) and/or front (32))** (Ex. 1005, 2:11-17), and physiological information from the **throat microphone (20)**. Ex. 1005, 2:31-53. The received signals are then compared with each other to distinguish speech as, e.g., voiced and unvoiced speech. Ex. 1005, 2:31-53. A POSITA would understand that Meisel describes this comparison process using cross correlation data between the acoustic information (e.g., side mouth microphone (18)) and the physiological information

(e.g., throat microphone (20)). Ex. 1003, ¶ 68. Meisel further discloses that the cross correlation data is sent to a data buffer (302), and then to a feature extractor (303) to examine a particular time period, such as sixty (60) milliseconds, following the first rise in the amplitude signal from the microphone (a portion of the acoustic signal received at the one receiver) “to determine if any of the logical signals 109, 110, 111 or 112 is true.” Ex. 1005, 13:5-13. If the feature extractor (303) recognizes “a voiced stop at the start of the word as indicated by a true signal on the line 112, a line 308 is set true.” Ex. 1005, 13:16-18. A POSITA would understand that classifying speech as a voiced stop involves determining whether the portion of the acoustic signal analyzed exceeds a correlation threshold. Ex. 1003, ¶ 68. As explained in § V.A.4.e ([1d]), Meisel’s signal pre-processor (2) compares the “m” sound portion of an acoustic signal from the nose microphone (16) with the “m” sound portion of an acoustic signal from the side mouth microphone (18) to detect a nasal speech. Ex. 1005, 8:24-59, FIGS. 11-12. A POSITA would understand from Meisel’s disclosure that the signals from the throat microphone (20) would be used to further classify the speech as voiced or unvoiced. Ex. 1005, 2:31-48; Ex. 1003, ¶ 68.

Moreover, identifying the acoustic signal as voiced speech when the **cross correlation data** corresponding to a portion of the acoustic signal received at one receiver **exceeds a correlation threshold** would have been obvious in view of

Freidin. Freidin discloses a correlator that correlates two received signals, an input data sequence and a reference sequence, to calculate a **correlation value y**. Ex. 1007, 1:43-46. A POSITA would recognize that correlation value y is cross correlation data because it is calculated “by first multiplying each data sample d_i by a corresponding reference value r_i ,” which is the same as the cross correlation equation that cross-correlates the input data samples to the reference values. Ex. 1007, 2:37-40. This **correlation value y** results in cross correlation data because it creates a correlation value for each data sample i. Ex. 1007, 1:14-22; Ex. 1003, ¶ 69.

Freidin further discloses that the “**correlation value y** [cross correlation data] is compared against a **correlation threshold y_{TH}** , and whenever y exceeds the **correlation threshold y_{TH}** , a correlation match is declared indicating that the correlation sequence has been received” Ex. 1007, 1:48-52.

A POSITA would have known that Freidin’s cross correlation calculation and comparison to a threshold value are applicable to Meisel’s speech recognition because Freidin discloses that “[c]orrelation is used in a number of applications including speech recognition.” Ex. 1007, 1:23-26. Freidin expressly states that the correlator operation “is not to be restricted except within the spirit and scope of the appended claims” because many variations of the disclosed embodiments “are possible which remain within the concept and scope of the invention,” which

“would become clear to one of ordinary skill in the art after inspection of the specification, drawings and claims.” Ex. 1007, 11:51-12:4. A POSITA would have found it obvious to apply Freidin’s cross correlator and threshold comparison to Meisel’s comparison of, for example, the side mouth microphone (18) and physiological throat microphone (20) because Meisel teaches correlating the received signals and using the result of the correlation to recognize speech. Ex. 1005, Abstract; Ex. 1003, ¶ 70. A POSITA would have found it obvious to use cross correlation because it was a conventional technique for comparing signals, which technique was a standard built-in function for signal processing software, such as MATLAB. Ex. 1003, ¶ 71. It also would have been obvious to compare the cross correlation data to a correlation threshold to provide an objective determination of the data to identify the desired speech type, such as voiced speech. Ex. 1005, 2:31-53; Ex. 1003, ¶ 71.

i. Motivation to combine

A POSITA would have been motivated to combine Meisel’s speech processing system with Freidin’s correlator to compare the correlation value (cross correlation data, *supra* § V.A.4.e) to a correlation threshold to identify the speech type as voiced speech in Meisel. Ex. 1003, ¶ 72. While Meisel and Freidin both correlate between signals, Freidin describes the cross correlation function to produce the cross correlation data and explains how to use that cross correlation

data to compare it with threshold y_{TH} to identify a result, such as Meisel's voiced speech. Ex. 1007, 1:48-52; Ex. 1003, ¶ 72.

A POSITA would have looked to Freidin to implement the cross correlation and speech identification performed by Meisel's ADCI and signal pre-processor (2) because Meisel describes a system where the signals are compared and correlated to determine a result, and Freidin provides a well-known cross-correlation-and-threshold comparison framework to render Meisel's identification more objective. Ex. 1003, ¶ 73. A POSITA would have been motivated to use the cross correlation function because it identifies voiced speech when there is a speech signal at both microphones (e.g., throat microphone and mouth microphone), and thus the difference (calculated by Meisel's ADCI) may be small. Ex. 1003, ¶ 73. Thus, because cross correlation allows for calculating the presence of voiced speech when the signal is present at both microphones (e.g., small difference, but high multiplication of signal values), it provides additional accuracy to Meisel's speech type determinations and expands to correlating other voiced speech types than just voiced stops. Ex. 1003, ¶ 73.

A POSITA would have been motivated to use Freidin's threshold comparison to identify voiced speech in Meisel because Freidin discloses that “[c]orrelators have been widely used in signal processing to measure the degree of

correspondence between a data sequence and a predetermined pattern or reference sequence,” e.g., in speech recognition. Ex. 1007, 1:7-26; Ex. 1003, ¶ 74.

ii. Expectation of success

A POSITA would have been motivated to combine Meisel and Freidin as discussed in § V.A.4.e above. It was also well known in the art to compare the cross correlation value to a predetermined threshold value to determine whether the signals sufficiently correlate with each other. Ex. 1007, 1:48-52; *e.g.*, Ex. 1014 (Chu), Abstract (“The signal is declared to be stationary in the current time interval when the cross correlation exceeds a threshold value, and is declared to be non-stationary in the current time interval when the cross-correlation is less than the threshold value.”). Accordingly, a POSITA would have reasonably expected to succeed in using Freidin’s cross correlation threshold comparison in Meisel’s speech recognition system to determine whether the signals represent a particular speech type, such as voiced speech, because the use of these techniques is nothing more than the application of conventional decades-old functions applied in their ordinary and expected manner. Ex. 1003, ¶ 75. Therefore, a POSITA would have found that implementing Freidin’s cross correlation threshold comparison in Meisel’s speech recognition system was combining prior art elements according to known methods (digital signal processing) to yield predictable results (improving the accuracy of speech detection). Ex. 1003, ¶ 75. Further, combining Meisel and

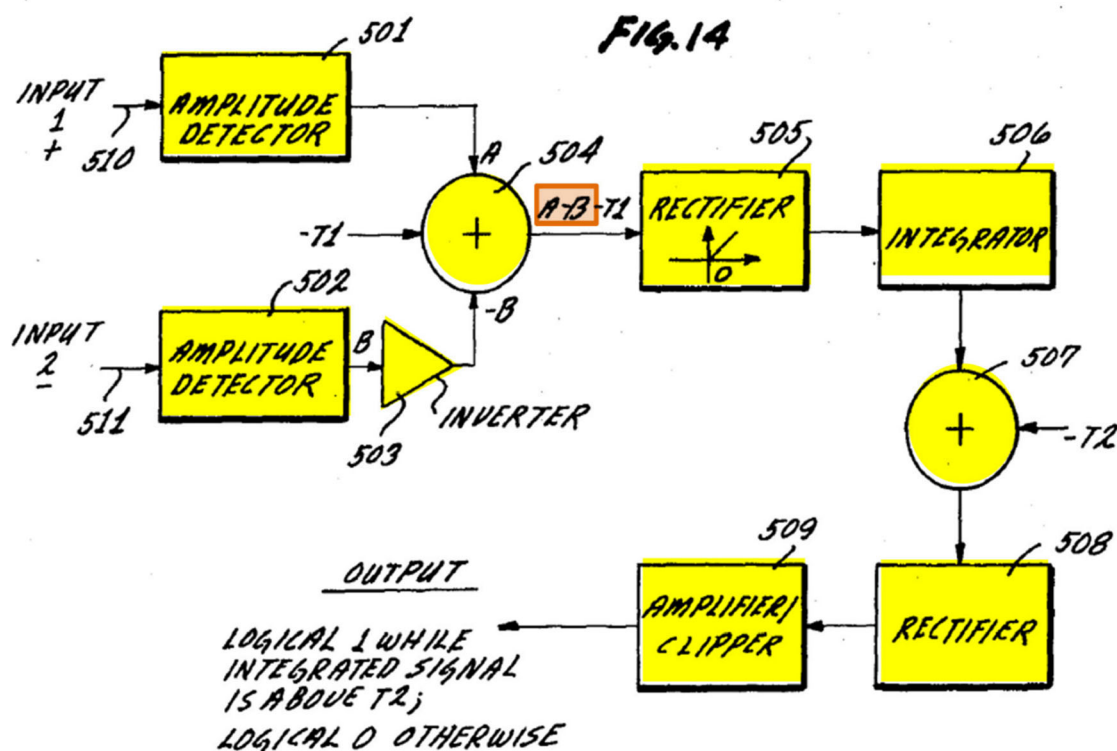
Freidin would have been applying a known technique (a correlator employing a correlation threshold) to a known device (a speech recognition system) ready for improvement to yield predictable results (improving the accuracy of speech detection).

- g. [1f] “generates difference parameters between the acoustic signals received at each of the two receivers, wherein the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals;”**

Meisel discloses that the processor (e.g., **signal pre-processor (2)**) generates difference parameters between the acoustic signals received at each of the two receivers (e.g., microphones 16, 18), wherein the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals. Ex. 1003, ¶ 76.

As explained in [1c] above (*supra* § V.A.4.d), different microphones in Meisel (e.g., **nose microphone (16)** and **mouth microphone (18)**) generate different signal amplitudes for the received signals, and the **signal pre-processor (2)** compares those signals to determine “instant imbalances in the relative amplitude of each microphone.” Ex. 1005, 4:35-39. Meisel’s “imbalances” refer to differences in the signal amplitudes received at different microphones, which a POSITA would recognize as the signal gain. Ex. 1003, ¶ 77. Meisel explains that “[t]hese imbalances are indicative of the production of certain speech sounds”

received at the microphones. Ex. 1005, 4:35-39. A POSITA would understand that the comparison of the signal amplitudes of the two microphones to detect instant imbalances, which a POSITA would understand determines or generates the difference (difference parameters) between the acoustic signals received at each of the two microphone receivers. Ex. 1003, ¶ 77. As shown below in Figure 14, the input signals (A, B) from two microphones are compared using the summing unit (504), which subtracts the amplitude of input signal B from the amplitude of input signal A, as shown by the “A-B” indicator below. Ex. 1005, 9:12-22.

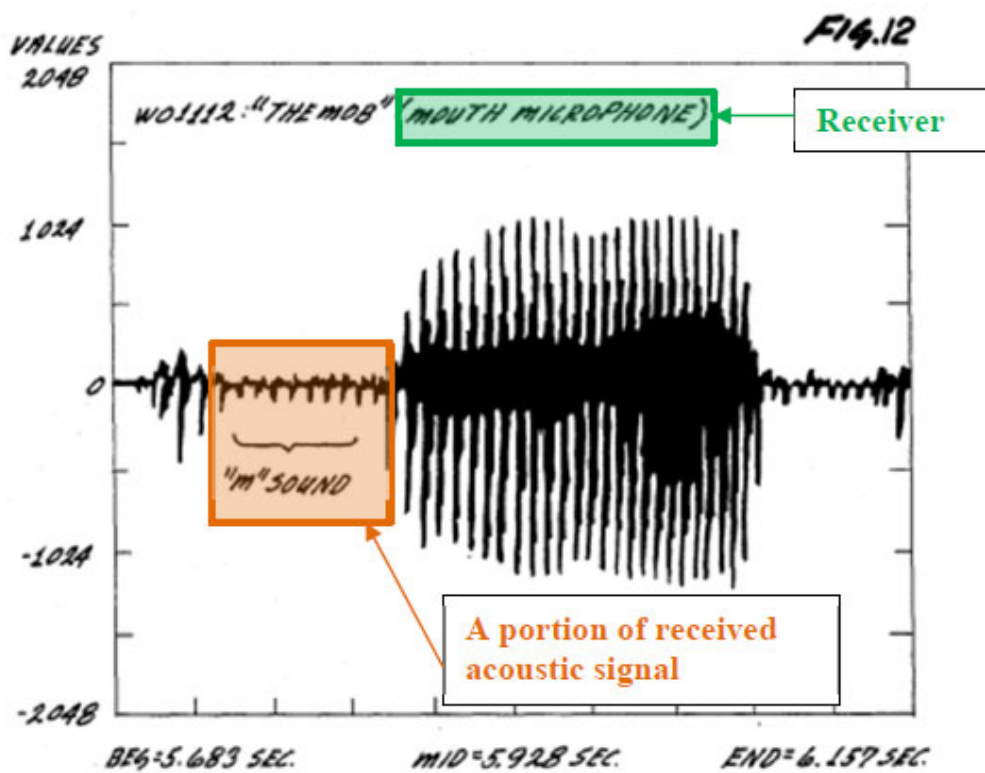
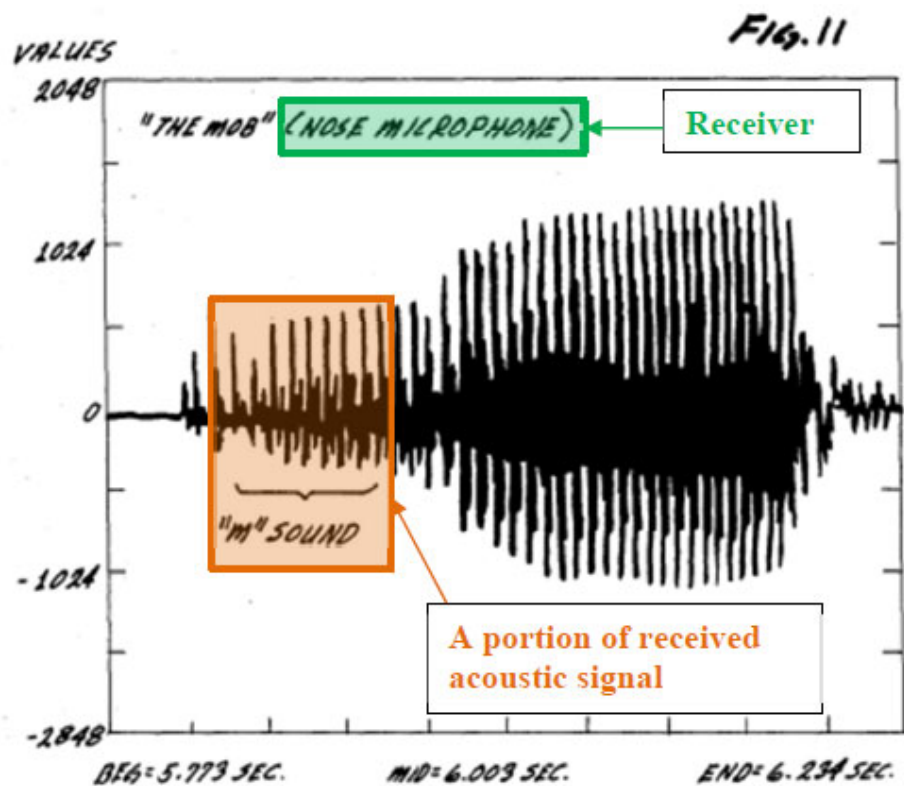


Ex. 1005, FIG. 14 (annotated).

As explained in § V.A.4.h ([1g]), *infra*, after the summing unit (504) calculates the difference between the amplitudes of input signals A and B, it compares the difference in the input signal amplitudes to a predetermined threshold T1. Ex. 1005, 9:12-22. If input signal A is greater than input signal B by at least T1, then the signals proceed to the next processing step to identify a speech type. Ex. 1005, 9:23-37. The integrator (506) processes the signals, and its output of the integrated amount (which measures by how much input signal amplitude A exceeds the input signal amplitude B) is fed to the comparator circuit (507) to compare the difference parameter (difference between input signal A and input signal B) to a second predetermined threshold T2. Ex. 1005, 9:38-42. The output of the comparator circuit (507) determines the type of speech, e.g., the occurrence of a nasal speech. Ex. 1005, 9:47-50; Ex. 1003, ¶ 78.

Meisel further discloses that the difference parameters are representative of the **relative difference in signal gain between portions of the received acoustic signals**. Ex. 1003, ¶ 79. Meisel explains that Figures 11 and 12 show the amplitude of acoustic signals from the nose microphone (16) and the first mouth microphone (18), respectively. Ex. 1005, 4:3-8. The values of the signals in Figures 11 and 12 represent the amplitude patterns recorded by the nose microphone (16) and the first mouth microphone (18) and are compared to determine whether the speech is nasal. Ex. 1005, 8:24-59; Ex. 1003, ¶ 79. A POSITA would understand that

amplitude signals are representative of the signal gain for each microphone, and the difference in gain is used to determine, e.g., voiced speech. Ex. 1003, ¶ 79.



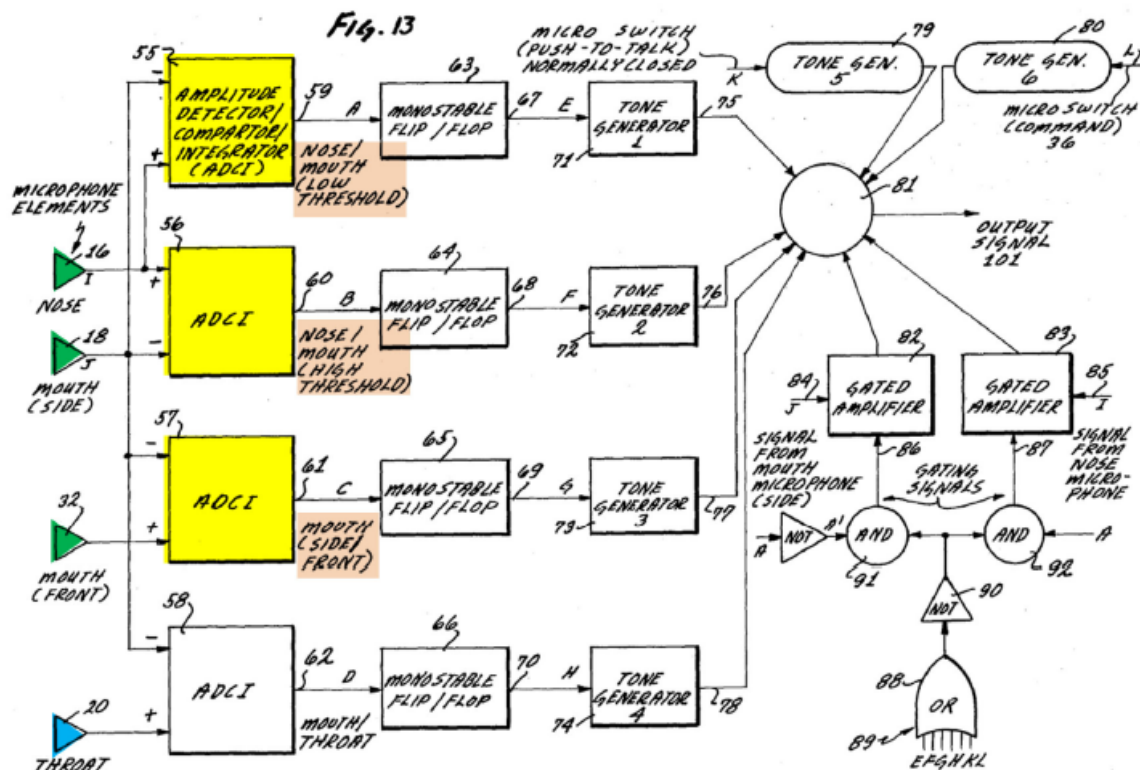
Ex. 1005, FIGS. 11-12 (annotated).

As one example in Meisel, the processor (**signal pre-processor (2)** and **speech recognition subsystem (3)**) recognizes the “M” sound in “The Mob” as a nasal speech because it recognizes that the signals received at the nose microphone (16) exhibit “a relatively large amplitude” during a period when the signals received at the first mouth microphone (18) exhibit “a relatively low amplitude.” Ex. 1005, 8:24-59; *see also* Ex. 1005, 2:25-30, 5:51-55, 10:1-3. A POSITA would understand that comparing a “large amplitude” to a “low amplitude” via subtraction is determining a difference in the signals received at the two microphones and constitutes generating the difference parameters between the nose microphone (16) and the first mouth microphone (18). Ex. 1003, ¶ 80. In other words, the relative difference in signal gain between the nose microphone (16) and the first mouth microphone (18) in the “M” portion of the acoustic signals allows the processor to determine the type of speech (e.g., nasal). Ex. 1003, ¶ 80. A POSITA would understand that the signal amplitude comparison in Figures 11 and 12 would be applied to the speech signals when a speaker’s mouth is closed, partially open, or open, each of which imparts different gains to the speech. Ex. 1003, ¶ 80; Ex. 1005, 9:66-10:7.

As explained further in [1g] below (*infra* § V.A.4.h), the summing unit (504), which acts as a comparator, generates difference parameters by inverting the amplitude of the second microphone signal (e.g., the nose microphone (16)), and

summing the inverted signal with the amplitude of the first acoustic signal. Ex. 1005, 9:12-22. Stated differently, the second acoustic signal amplitude is subtracted from the first acoustic signal amplitude to generate the difference parameter between the two microphones. Ex. 1005, 9:12-22; Ex. 1003, ¶ 81.

A POSITA would understand that the same process would be applied to other microphones, e.g., between the first mouth microphone (18) at the corner of the speaker's mouth and the second mouth microphone (32) in front of the speaker's mouth. Ex. 1003, ¶ 82. In fact, Meisel expressly states that “[s]imilar proof can be provided of the importance of recording the sounds emanating from the throat or immediately in front of the mouth for certain words in order to provide a reliable recognition of such spoken words.” Ex. 1005, 8:55-59. Based on this disclosure, a POSITA would understand that the acoustic signals from the first mouth microphone (18) can be compared with those of the second mouth microphone (32) to determine a difference in their signal gains and generate difference parameters between the two mouth microphones. Ex. 1003, ¶ 82.



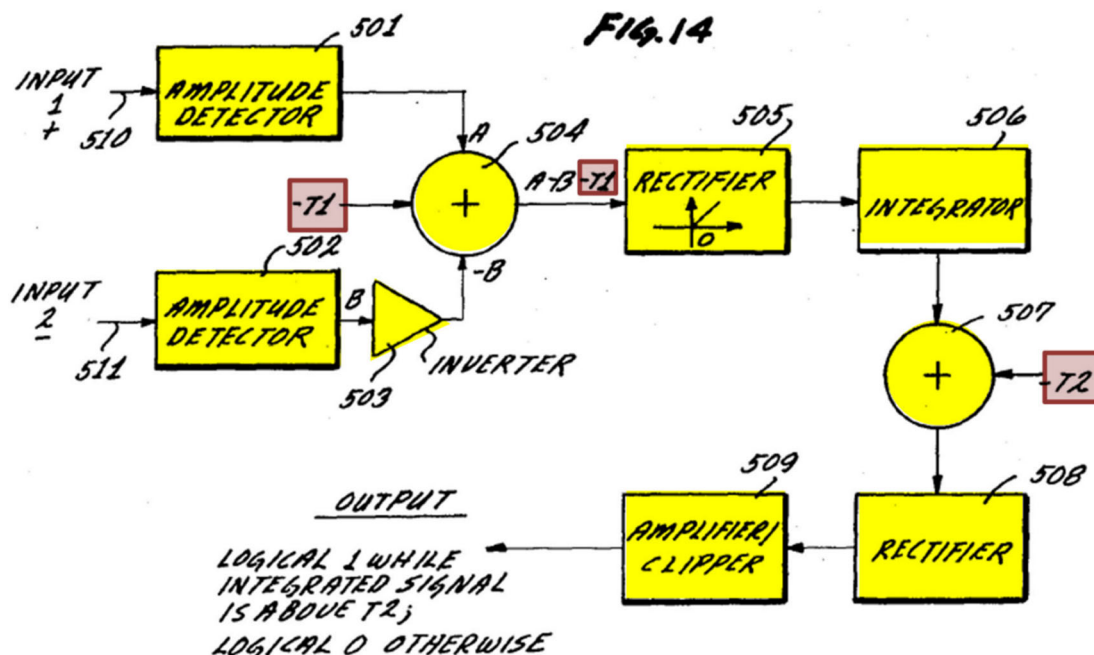
Ex. 1005, FIG. 13 (annotated).

- h. [1g] “identifies information of the acoustic signals as unvoiced speech when the difference parameters exceed a gain threshold; and”

Meisel discloses that the processor (**signal pre-processor (2)**) identifies information of the acoustic signals as unvoiced speech when the difference parameters exceed a **gain threshold**. Ex. 1003, ¶ 83.

Meisel discloses that the acoustic signals received at a pair of microphones in the microphone assembly (comprising the nose microphone (16) and two mouth microphones (18, 32)) are input into the ADCI units. The ADCI units are part of

the signal pre-processor (2) and include two gain thresholds, T1 and T2, to which the differences are compared. Ex. 1005, 8:60-9:2, 9:12-42; Ex. 1003, ¶ 84.

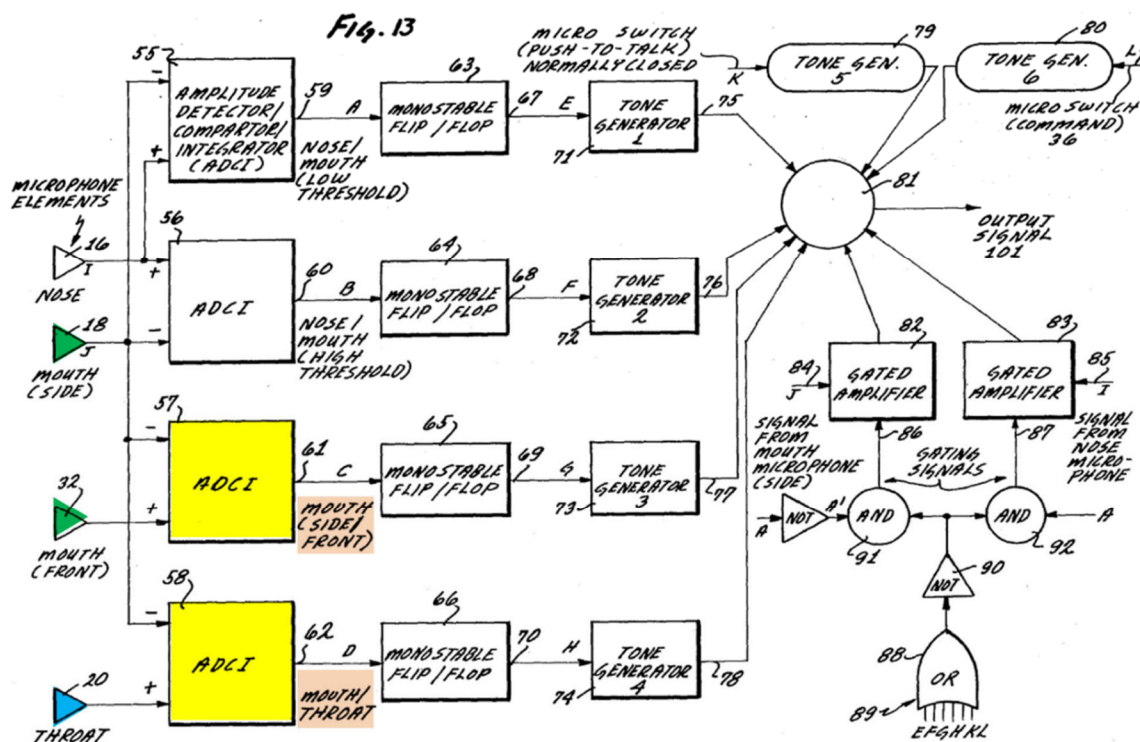


Ex. 1005, FIG. 14 (annotated).

As explained in [1e] above, Meisel discloses the feature extractor (303), which distinguishes voiced and unvoiced stop consonants. *Supra* § V.A.4.f. Meisel explains that this is accomplished when each of the amplitude detectors (501, 502) receives inputs (acoustic signals, the first input and the second input) from a pair of microphones. Ex. 1005, 8:63-9:11 The amplitude detectors (510, 511) determine the amplitudes of the input signals. Ex. 1005, 9:4-11. The inverter (503) inverts the signal amplitude (Input 2) from the second microphone (e.g., nose microphone

(16) or second mouth microphone (32)), such that it becomes a negative value (-B).

Ex. 1005, 9:12-14. The summing unit (504) adds the signal amplitude of the first input and inverted signal amplitude of the second input, and therefore generates difference parameters between the two signals. Ex. 1005, 9:14-20; Ex. 1003, ¶ 85



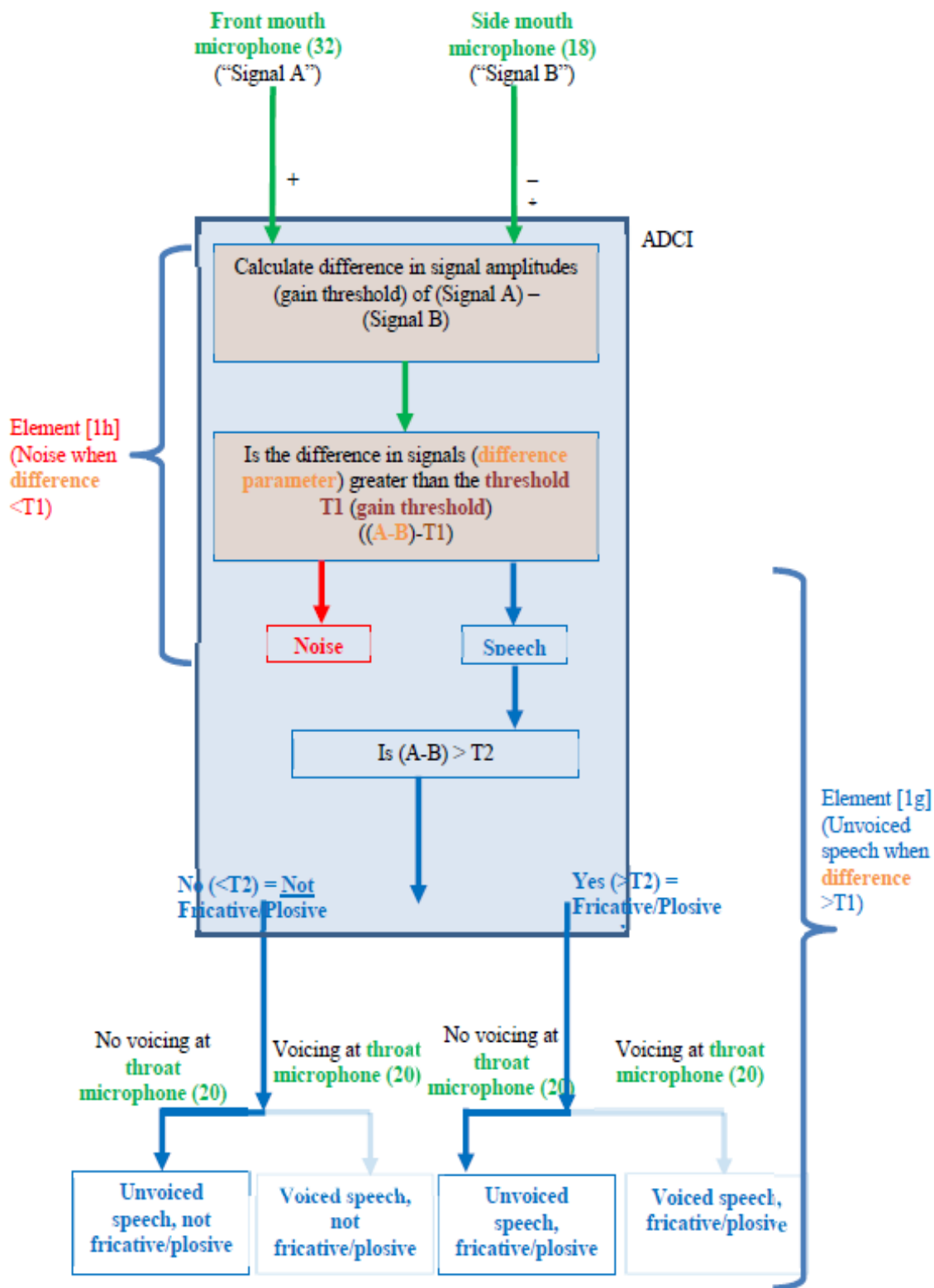
Ex. 1005, FIG. 13 (annotated).

Meisel discloses that this signal comparison step would be applied to the side mouth microphone (18) and the front mouth microphone (32) to determine unvoiced speech and noise. Ex. 1005, 8:60-9:50. The signal inputs from both microphones are fed to the summing unit (504) of ADCI unit (57), which compares

the amplitudes of the signals and determines whether the difference between the two signals is above a predetermined threshold T1 (a **gain threshold**). Ex. 1005, 9:12-37. The output of the summing unit (504) is fed to the rectifier (505), which ensures “that there is no output from the rectifier 505 unless the signal at the input is positive.” Ex. 1005, 9:23-25. A positive signal input to the rectifier (505) happens only when the first input (e.g., front mouth microphone (32)) is greater than the second input (e.g., side mouth microphone (18)), and the difference between the first input and the second input (A-B) exceeds the predetermined threshold (T1). Ex. 1005, 9:23-37. In view of Figure 13, a POSITA would understand that the output (61) would also be correlated or compared with the signals from the throat microphone (20) to distinguish the speech as voiced and unvoiced because the throat microphone (20) indicates when speech is voiced or not. Ex. 1003, ¶ 86; Ex. 1005, 2:31-48.

If the difference between the signal inputs from the front mouth microphone (32) and the side mouth microphone (18) is greater than the **gain threshold T1**, Meisel’s system proceeds to compare the difference (A-B) to a second threshold to determine the type of speech, for example, whether the speech is a fricative or a plosive. Ex. 1005, 9:38-50, 10:7-9. T1 in Meisel is a gain threshold because it is a threshold for acoustic signal amplitudes designed to distinguish voiced and unvoiced speech. Ex. 1003, ¶ 87. A POSITA would have recognized that Meisel’s

disclosure of the throat microphone (20) would also increase the accuracy of speech recognition when combined or compared with various speech types (e.g., fricative/plosive identification) to distinguish voiced and unvoiced speech. Ex. 1005, 10:9-11; Ex. 1003, ¶ 87. A POSITA would have found it obvious to combine the output (61) of ADCI unit (57) with the cross correlation data explained in § V.A.4.e ([1d]) above to distinguish voiced speech from unvoiced speech. Ex. 1003, ¶ 87. A POSITA would also understand that a positive (non-zero) output of the summing unit (504) in ADCI unit (57) denotes the difference parameters exceeding a **gain threshold (T1)**. Ex. 1003, ¶ 87. This process is summarized below in Flow Chart 1.



Flow Chart 1.

A POSITA would recognize that Meisel performs generating the difference parameters and comparing the difference parameters to a gain threshold in two steps. Ex. 1003, ¶ 88. Meisel discloses that signal amplitude A is summed with the inverted signal amplitude B and compared to the gain threshold T1 (Ex. 1005, 9:14-20), which Figure 14 describes as “A-B-T1” (Ex. 1005, FIG. 14). A POSITA would understand that, based on the simple order of operation, A-B must precede -T1 as a separate step, and these steps occur separately. Ex. 1003, ¶ 88. Such multiple subtractions were conventionally resolved through multiple steps, rather than a single double subtraction. Ex. 1003, ¶ 88. Furthermore, a POSITA would recognize that it was less expensive and easier to calculate the A-B and -T1 steps separately when using logic circuits and also negligible in processing time when using digital signal processing. Ex. 1003, ¶ 88. Therefore, a POSITA would understand that Meisel discloses generating difference parameters (A-B) as a separate step from comparing the difference parameter to the gain threshold (-T1). Ex. 1003, ¶ 88.

- i. **[1h] “identifies information of the acoustic signals as noise when the difference parameters are less than the gain threshold.”**

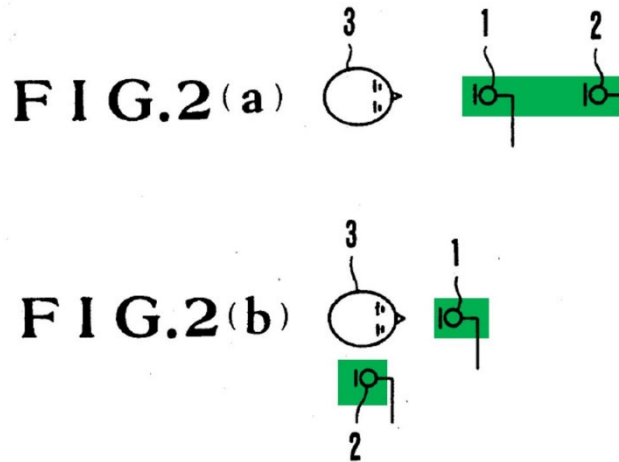
Meisel discloses that its speech detection system operates when there is background noise or other types of noise associated with the speaker performing other duties, suggesting that the noise is detected to minimize its effects in accurate

speech recognition. Ex. 1005, 1:10-15 (“More particularly, the invention relates to a speech recognition system usable in the normal work environment of the speaker, particularly under conditions where there is background noise and where the speaker can continue to perform other duties.”). Meisel cites Sakoe, which teaches how “to detect background noise with a view toward minimizing its effect on recognition accuracy.” Ex. 1005, 1:46-50; Ex. 1003, ¶ 89.

Meisel discloses that its comparison process illustrated in Figure 13 above, *supra* § V.A.4.h ([1g]), distinguishes speech from noise. Ex. 1003, ¶ 90. When the difference between the first mouth microphone input and the second mouth microphone input (A-B) is below the threshold T1, then the information of the acoustic signals from the side and front mouth microphones (18, 32) is identified as noise. Ex. 1003, ¶ 90; *supra* § V.A.4.h (Flow Chart 1).

Moreover, “identif[ying] information of the acoustic signals as noise when the difference parameters are less than the gain threshold” would have been obvious to a POSITA in view of Kaneda. Kaneda discloses a “conventional” speech recognition system comprising **two acoustic microphones** where the microphones are located near a speaker (microphone 1) and away from the speaker (microphone 2), respectively. Ex. 1006, 1:51-57, FIG. 2(a). During a speech period, the speech signals at microphone 1 have a higher power than those received

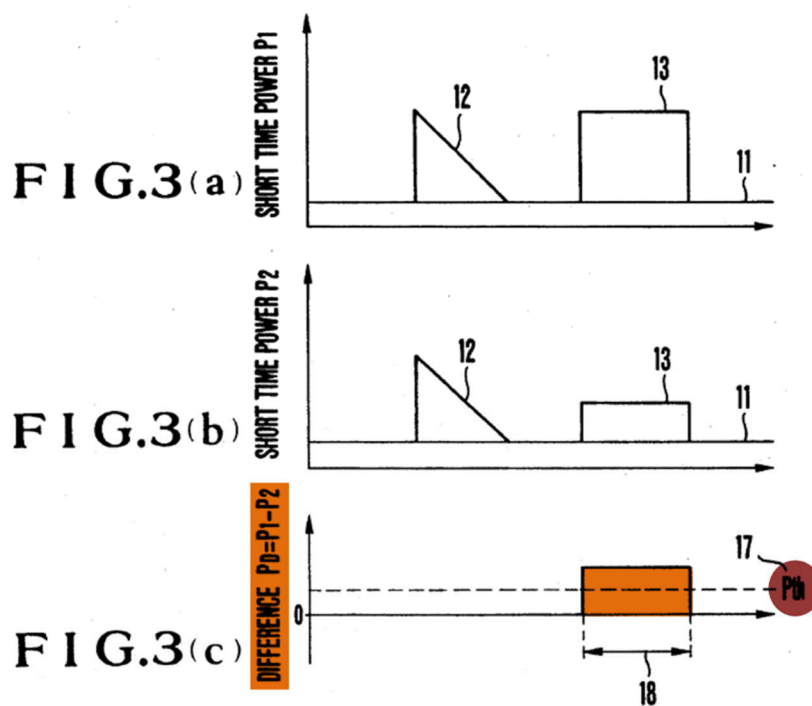
at microphone 2 because microphone 1 is closer to the speaker. *Id.*, 1:61-63;
compare Figure 3(a) *with* Figure 3(b); Ex. 1003, ¶ 91.



Id., FIGS. 2(a)-2(b) (annotated).

This “conventional” speech recognition system calculates a difference, PD, between the speech power outputs of the two microphones “PD (=P1-P2)” (**difference parameters**), and when PD is greater than **threshold value Pth17**, “a corresponding time period 18 is detected as a speech period.” *Id.*, 2:12-18.

Threshold value Pth17 is a **gain threshold** because it is a signal amplitude threshold used to discriminate speech and noise. If the **power difference PD** is less than **Pth17**, then the system identifies the acoustic signals as noise. Ex. 1003, ¶ 92, Ex. 1006, 2:18-22.



Ex. 1006, FIGS. 3(a)-3(c) (annotated).

i. Motivation to combine

A POSITA would have been motivated to combine Meisel's speech processing system with Kaneda's conventional speech and noise determination using difference parameters between two acoustic input signals. Ex. 1006, 1:51-2:22; Ex. 1003, ¶ 93-94. Kaneda, like Meisel, is directed to a speech recognition system using two or more acoustic microphones. *Supra* §§ V.A.1-V.A.2, V.A.4.b. Meisel discloses that the acoustic signals from two acoustic microphones would be compared with each other to determine the difference in their signal amplitudes. Ex. 1005, 4:35-39, 8:24-59, FIGS. 11-12. Kaneda, like Meisel, uses a difference in signal amplitudes, and explains how the difference would be used to distinguish

between noise and speech instead of determining speech type. *Compare* Ex. 1005, 9:12-50, *with* Ex. 1006, 2:12-22.

While Meisel does not expressly disclose identifying acoustic signals as noise, Meisel discloses that its speech recognition system is “usable in the normal work environment of the speaker, particularly under conditions where there is background noise and where the speaker can continue to perform other duties.” Ex. 1005, 1:10-15. Furthermore, a POSITA would recognize that Kaneda’s teachings are directly implementable in Meisel’s system because both involve finding signal amplitudes, finding the differences in those amplitudes, and comparing the differences to a threshold. Ex. 1003, ¶ 95. These are all functions of Meisel’s ADCI units. Ex. 1003, ¶ 95. Incorporating the results of Kaneda’s conventional speech detection into Meisel’s system would be trivial to a POSITA because Meisel already discloses a set of signals that conveys the detected characteristics of the speech to a processor for speech recognition. Ex. 1003, ¶ 95. A POSITA would have been motivated to look to Kaneda’s noise determination to remove noise from Meisel’s acoustic detection to better identify speech signals in Meisel’s speech recognition. Ex. 1003, ¶ 95. Therefore, a POSITA would have looked to Kaneda to identify noise periods to mitigate errors in speech recognition. Ex. 1003, ¶ 95.

Although Kaneda discloses three conditions that must be present for the conventional speech recognition system to accurately distinguish speech and noise

in practical applications, Ex. 1006, 2:23-27, a POSITA would recognize that Meisel meets all three conditions. Ex. 1003, ¶ 96. First, Kaneda discloses that “[a]n S/N ratio difference in two signals must be present,” which Kaneda acknowledges is satisfied in conventional speech recognition systems, such as Meisel’s. Ex. 1006, 2:28-29. Indeed, as explained in Sections V.A.2 (Overview of Meisel) and V.A.4.b (element [1a]), Meisel’s speech recognition system includes at least two acoustic microphones (a nose microphone (16), a first mouth microphone (18), and a second mouth microphone (32)), which provide different acoustic signal strengths that Meisel uses to calculate the power differences in the ADCIs. Ex. 1005, 2:11-30, 8:37-59, FIGS. 11-12. Second, Kaneda discloses that “[n]oise and speech periods of the two signals must be matched with each other as a function of time.” Ex. 1006, 2:30-32. A POSITA would understand that Meisel’s system satisfies this condition because its microphones are located close to each other in a small, unitary multi-microphone structure. Ex. 1003, ¶96 Ex. 1005, FIG. 5. Due to the close proximity of the microphones, any temporal difference in noise and speech periods of the two signals would be minimal and negligible. Ex. 1003, ¶ 96. Third, Kaneda discloses that “[a] variation in S/N ratio difference caused by various factors is small (stability of the S/N ratio difference).” Ex. 1006, 2:33-35. Given the close proximity between the microphones in Meisel’s system, a POSITA would understand that the noise levels received at the microphones would be similar,

leading to a stable speech to noise (S/N) ratio difference between the microphones. Ex. 1003, ¶ 96. A POSITA would therefore have recognized that Meisel's system was appropriate for using Kaneda's "conventional" speech detection method of $PD = P1 - P2$, where PD is compared to a threshold value to differentiate speech from noise. Ex. 1003, ¶ 96.

While Kaneda discloses that "variation factors"³ in conventional speech recognition systems may cause an unstable S/N ratio, Ex. 1006, 3:13-48, a POSITA would have recognized that the effects of the variation factors are minimal in Meisel's system because of the close proximity of the microphones to each other. Ex. 1003, ¶ 97; Ex. 1005, FIG. 5. Indeed, a POSITA would have been motivated to implement Kaneda's conventional speech detection method in Meisel's speech recognition system to minimize the effects of the variation factors in accurate speech recognition. Ex. 1003, ¶ 97.

Although Kaneda acknowledges that, in some circumstances, "the second conventional method has [a] decisive drawback," referring to the variation factors discussed above (Ex. 1006, 3:49-51), a POSITA would understand that this drawback does not apply to Meisel's speech recognition system as explained

³ These factors include the relative position of the noise source, movement of the speaker, and room echoes, which vary as a function of time. Ex. 1006, 2:13-48.

above. Ex. 1003, ¶ 98. Meisel's microphones are located very close to each other, making the effects of the variation factors trivial among the microphones. Meisel's closely spaced microphone arrangement, therefore, meets the conditions for the conventional system in Kaneda to operate accurately. Ex. 1003, ¶ 98; *compare* Ex. 1005, FIG. 5 *with* Ex. 1006, FIGS. 6(a)-6(b).

ii. Expectation of success

A POSITA would have reasonably expected to succeed in combining Kaneda's acoustic signal processing to identify noise by generating difference parameters from acoustic microphones with Meisel's system. Ex. 1003, ¶ 99.

Meisel describes a speech recognition system that includes three acoustic microphones (a nose microphone (16) and two mouth microphones (18, 32)). Ex. 1005, 2:11-17, 2:31-33. Kaneda's conventional speech recognition system similarly includes a microphone array having two acoustic microphones. Ex. 1006, 1:51-2:22. Also, Kaneda, like Meisel, discloses comparing acoustic signal amplitudes from the microphones to calculate differences in received signal amplitudes. Ex. 1006, 2:12-18. Given the similarities in the system architecture and the general approach of comparing signal amplitudes, a POSITA would have had a reasonable expectation of success in combining Meisel's speech recognition system with Kaneda's conventional signal processing technique of generating difference parameters. Ex. 1003, ¶ 100. Kaneda's conventional signal processing

for detecting noise uses the signals received and difference parameters already calculated in Meisel's system, adding an additional signal processing step to detect noise. Therefore, implementing Kaneda's noise detection in Meisel's system would not inhibit Meisel's signal processing functions, which are used to determine the types of speech for speech recognition. Ex. 1003, ¶ 100. For example, Kaneda's conventional noise detection would be implemented in parallel with Meisel's speech detection. Ex. 1003, ¶ 100.

Adding Kaneda's conventional signal processing technique for detecting noise using difference parameters merely combines known prior art elements according to known and expected functions, such as a known speech signal processing technique to calculate a difference between acoustic signals by subtracting one signal amplitude from another and comparing the difference to a predetermined threshold. Ex. 1003, ¶ 101. This combination would yield predictable results to accurately determine the type of speech or noise. Ex. 1003, ¶ 101. Also, it would have been applying a known technique (a known speech signal processing technique of calculating a difference between acoustic signals by subtracting one signal amplitude from another and comparing the difference to a predetermined threshold) to a known device ready for improvement (a speech recognition system) to yield predictable results (accurately determine the type of speech). Ex. 1003, ¶ 101.

5. Independent Claim 2

Meisel teaches or suggests each element of claim 2.

a. [2Preamble] “A method for removing noise from acoustic signals, comprising:”

As explained for [1h] (*supra* § V.A.4.i), Meisel discloses that detecting noise in acoustic signals to minimize its effects in speech detection was well known in the art. Ex. 1003, ¶ 103. Moreover, this element would have been obvious in view of Kaneda. *Id.*

Notably, neither “removing noise” nor “denoising” is a recited step of independent claim 2. It is, instead, recited in dependent claim 4, discussed in § V.A.6, *infra*. Thus, removing noise is not a required element of claim 2. Removing noise would have been obvious as discussed for dependent claim 4. Ex. 1003, ¶ 104.

A POSITA would understand that identifying noise disclosed in Meisel would be beneficial to suppress or mitigate the impact of noise on speech recognition. Ex. 1003, ¶ 105. Removing noise would therefore have been obvious to provide better speech recognition in environments where noise exists. *Id.* Kaneda explains that filter control for minimizing the power of a noise component is a “well-known minimization problem” in speech recognition. Ex. 1006, 8:29-32. As such, a POSITA would have been motivated to mitigate this problem using denoising methods known in the art. Ex. 1003, ¶ 105.

For example, Sakoe—cited in the background section of Meisel—describes one such noise-removing method. Meisel cites Sakoe and explains that Sakoe’s system “detect[s] background noise with a view toward minimizing its effect on recognition accuracy.” Ex. 1005, 1:46-50. Sakoe discloses a speech recognition system, which includes a “rejection unit which, when the intensity of said noise exceeding a predetermined standard value is detected in any of said speech signal durations, rejects the recognition results determined by said recognition unit.” Ex. 1008, 2:17-33. A POSITA would understand that rejecting the recognition results is removing noise from acoustic signals. Ex. 1003, ¶ 106.

b. [2a] “receiving the acoustic signals at two receivers and receiving physiological information associated with human voicing activity at a voicing sensor;”

[2a] is substantially similar to [1a]-[1c], and is disclosed or rendered obvious as discussed for [1a]-[1c]. *Supra* §§ V.A.4.b-V.A.4.d; Ex. 1003, ¶ 107.

c. [2b] “generating cross correlation data between the physiological information and an acoustic signal received at one of the two receivers;”

[2b] is identical to [1d], and is disclosed or rendered obvious as discussed for [1d]. *Supra* § V.A.4.e; Ex. 1003, ¶ 108.

d. [2c] “identifying information of the acoustic signals as voiced speech when the cross correlation data corresponding to a portion of the acoustic signal received at the one receiver exceeds a correlation threshold;”

[2c] is substantially identical to [1e], and is disclosed or rendered obvious as discussed for [1e]. *Supra* § V.A.4.f; Ex. 1003, ¶ 109.

- e. **[2d] “generating difference parameters between the acoustic signals received at each of the two receivers, wherein the difference parameters are representative of the relative difference in signal gain between portions of the received acoustic signals;”**

[2d] is substantially identical to [1f], and is disclosed or rendered obvious as discussed for [1f]. *Supra* § V.A.4.g; Ex. 1003, ¶ 110.

- f. **[2e] “identifying information of the acoustic signals as unvoiced speech when the difference parameters exceed a gain threshold; and”**

[2e] is substantially identical to [1g], and is disclosed or rendered obvious as discussed for [1g]. *Supra* § V.A.4.h; Ex. 1003, ¶ 111.

- g. **[2f] “identifying information of the acoustic signals as noise when the difference parameters are less than the gain threshold.”**

[2f] is substantially identical to [1h], and is disclosed or rendered obvious as discussed for [1h]. *Supra* § V.A.4.i; Ex. 1003, ¶ 112.

6. Claim 4: “further comprising performing denoising on the identified noise.”

Claim 4 depends from claim 2. Meisel and Kaneda render obvious claim 4.

Meisel discloses performing denoising on the identified noise. As explained in [1h], Meisel and Kaneda disclose a process of identifying noise. *Supra* §

V.A.4.i. A POSITA would understand that the identification and minimization of noise disclosed in Meisel and Kaneda correspond to “performing denoising on the identified noise.” Ex. 1003, ¶ 114. As explained in [2Preamble] (*supra* § V.A.5.a), Sakoe further discloses a noise removal process, indicating that denoising or noise removal was well known at the time of the invention. *Id.*

Sakoe, like Meisel and Kaneda, is directed to “[a] speech recognition system adaptable to noisy environments,” which includes “a recognition unit for recognizing input speech signals and a noise measuring unit for measuring the intensity of ambient noises,” and “a rejection unit responsive to a rejection standard controlled by the intensity of the measured noise” Ex. 1008, Abstract. Therefore, a POSITA would have been motivated to combine Sakoe’s noise removal process with the speech recognition systems disclosed in Meisel and Kaneda. Ex. 1003, ¶ 115. Kaneda also discloses detecting noise and minimizing its effect on speech recognition. Ex. 1006, 4:59-63. Therefore, a POSITA would have had a reasonable expectation of success in combining Sakoe with Meisel and/or Kaneda. *Id.*

7. **Claim 5: “wherein the voicing sensor includes at least one detector selected from a group including radio frequency devices, electroglottographs, ultrasound devices, acoustic throat microphones, and airflow detectors.”**

Claim 5 depends from claim 2. Meisel renders claim 5 obvious.

As explained in [1b] (*supra* § V.A.4.c), Meisel discloses using an acoustic “throat microphone” (20) to “detect[] the vibration of the vocal cords, or ‘voicing.’” Ex. 1005, 2:31-35. A POSITA would understand that Meisel’s throat microphone (20) is an acoustic throat microphone because it “is disposed to detect vibration of the vocal cords through sounds emanating from the throat of a speaker.” Ex. 1005, 5:29-36; Ex. 1003, ¶ 116.

B. [Ground 2] Claim 3 Is Rendered Obvious by Meisel in View of Kaneda, Freidin, and Lyons

1. Lyons

Lyons discloses averaging digital signals using “statistical measures known as the mean, variance, and standard deviation” to minimize effects from unwanted random disturbances (noise) when measuring a parameter. Ex. 1012, 319, 477.

Lyons teaches that the standard deviation of digital signals is “defined as the positive square root of the variance, or

$$\sigma = \sqrt{\frac{1}{N} \sum_{n=1}^N [x(n) - x_{ave}]^2} . \quad (8-3).”$$

Ex. 1012, 320.

“If the sequence $x(n)$ represents a time series of signal samples, we can say that x_{ave} specifies the constant, or DC, value of the signal, [and] the standard

deviation σ reflects the amount of the fluctuating, or AC, component of the signal” *Id.*, 320, 479.

Lyons’s digital signal processing has many applications, including “voice recognition[and] data communication” (*id.*, 1), and cites to 22 publications directed to digital signal processing used in speech recognition (*id.*, 21, 47, 127, 128, 155, 156, 217, 218, 295, 317, 318, 441, 442). A POSITA would understand that the standard deviation disclosed in Lyons is applicable to the speech recognition processes disclosed in Meisel. Ex. 1003, ¶ 120.

Lyons was publicly available as 102(b) prior art because it was catalogued, indexed, and shelved in the Library of Congress on or about June 24, 1996, and the first published on October 31, 1996, as shown by the Copyright Office’s Copyright registration certificate. Exs. 1027-1028; Ex. 1003, ¶ 121.

2. Claim 3: “further comprising generating the gain threshold using standard deviations corresponding to the generation of the difference parameters.”

Claim 3 depends from claim 2. Lyons confirms and renders obvious using standard deviations corresponding to the generation of the difference parameters, as recited in claim 3.⁴

⁴ Claim 3 appears to lack written description and be indefinite as written. However,

Meisel discloses the gain threshold as explained in [1g]. *Supra* § V.A.4.h.

Meisel does not expressly disclose generating the gain threshold using standard deviations corresponding to the generation of the difference parameters, but it would have been obvious to a POSITA based on well-known statistical techniques in the art. A POSITA would understand that applying this well-known statistical technique to calculate standard deviations of the differences in the received signals (alternatively, the difference of the standard deviations of each received signal) represents the general variation that corresponds to the meaningful threshold difference in the signals (gain thresholds). Ex. 1003, ¶ 123; Ex. 1001, 8:41-47.

Indeed, it was well known at the time that calculating a standard deviation is a common statistical technique used for digital signal processing. Ex. 1012, 479 (explaining that a “useful measure of a signal sequence is the square root of the variance known as the *standard deviation*”); Ex. 1011, 14 (explaining that for most acquired signals that “do not show a well defined peak-to-peak value, but have a random nature ... [, a] more generalized method must be used in these cases, called the **standard deviation**, denoted by σ (a lower case Greek *sigma*)”). Lyons discloses that the standard deviation of digital signals is “defined as the positive square root of the variance, or

Petitioner has attempted to apply its best understanding of this term as written.

$$\sigma = \sqrt{\frac{1}{N} \sum_{n=1}^N [x(n) - x_{ave}]^2} . \quad (8-3).''$$

Ex. 1012, 320.

Moreover, a computer algorithm for calculating standard deviation in speech signal processing was also well known at the time. Ex. 1003, ¶ 125. A POSITA would recognize that applying standard deviations to generating the gain threshold for speech signal processing would have been obvious because the standard deviation represents the background power in signal processing that corresponds to the threshold at which the difference in signals is meaningful. Ex. 1003, ¶ 125.

i. Motivation to combine

A POSITA would have been motivated to apply standard deviation to Meisel's speech recognition system as modified by Kaneda and Freidin, based on what was known in the art at the time. Ex. 1003, ¶ 126. It was well known at the time that standard deviation is a common statistical measure to calculate an average deviation of input digital signals and minimize the effects of noise in measuring the true value of a parameter. Ex. 1012, 319, 476. A POSITA would understand that the standard deviation is applicable to the processing of speech signals disclosed in Meisel and Kaneda because they are a type of digital signal. Ex. 1003, ¶ 126. Moreover, Lyons cites to 22 different publications related to speech signal processing, demonstrating that the digital processing techniques

disclosed in Lyons, including the standard deviation calculation, are applicable to speech processing. Ex. 1012, 21, 47, 127, 128, 155, 156, 217, 218, 295, 317, 318, 441, 442. A POSITA would understand that speech recognition techniques are derived from digital signal processing techniques disclosed in Lyons. Ex. 1003, ¶ 126.

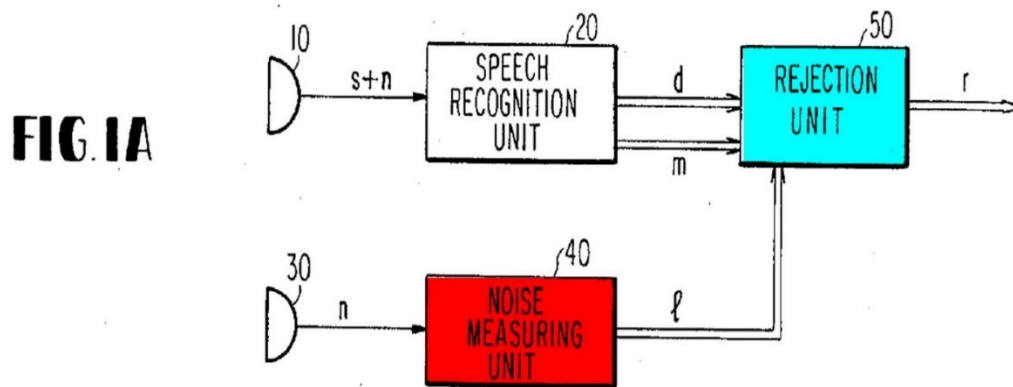
ii. Expectation of success

A POSITA would have had reasonable expectation of success in applying standard deviation to the speech recognition processes disclosed in Meisel and Kaneda as modified by Freidin because doing so would have been combining prior art elements (difference parameters and a standard deviation) according to known methods (using any one of known methods of calculating standard deviations for speech signals) to yield predictable results (obtaining a standard deviation of the speech signals to generate a gain threshold). Ex. 1003, ¶ 127. Also, it would have been obvious to try because standard deviation is only one of three statistical methods (mean, variance, and standard deviation) for speech signal processing. Ex. 1012, 477. Therefore, generating a gain threshold using standard deviations corresponding to the generation of the difference parameters would have been choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success. Ex. 1003, ¶ 127.

C. [Ground 3] Claim 4 Is Rendered Obvious by Meisel in View of Kaneda, Freidin, and Sakoe

As explained in § V.A.6, Meisel, Kaneda, and Freidin render claim 4 obvious. Claim 4 is also rendered obvious by Meisel in view of Kaneda, Freidin, and Sakoe.

Like Meisel and Kaneda, Sakoe is directed to “[a] speech recognition system adaptable to noisy environments,” which includes “a recognition unit for recognizing input speech signals and a noise measuring unit for measuring the intensity of ambient noises,” and “a rejection unit responsive to a rejection standard controlled by the intensity of the measured noise” Ex. 1008, Abstract. Therefore, Sakoe discloses “performing **denoising** on the **identified noise**,” and when combined with Meisel, Kaneda, and Freidin, renders claim 4 obvious.



Ex. 1008, FIG. 1A (annotated).

i. Motivation to combine

A POSITA would have been motivated to combine Sakoe's noise removal process with the speech recognition systems disclosed in Meisel and Kaneda. Ex. 1003, ¶ 131. As explained in § V.A.4.i ([1h]), *supra*, Meisel discloses a process of detecting speech signals. Meisel's background section explains that Sakoe "detect[s] background noise with a view toward minimizing its effect on recognition accuracy." Ex. 1005, 1:46-50. Like Sakoe, Kaneda discloses detecting and minimizing noise effects on speech recognition. Ex. 1006, 4:59-63. Therefore, both Meisel and Kaneda, as modified by Freidin, provide some teaching, suggestion, or motivation to remove noise to improve the accuracy of Meisel's speech recognition system. Ex. 1003, ¶ 131.

ii. Expectation of success

A POSITA would have reasonably expected to succeed in combining Sakoe's denoising process to remove noise. Ex. 1003, ¶ 132. Combining Sakoe with Meisel and Kaneda, as modified by Freidin, would have been combining prior art elements (speech recognition system and denoising system) according to known methods (known denoising process) to yield predictable results (remove noise to increase the accuracy of speech detection). A POSITA also would have expected the use of a known technique (denoising) to improve similar devices (speech recognition systems) to obtain predictable results (removing noise to increase the accuracy of speech detection). *Id.*

VI. DISCRETIONARY DENIAL UNDER §314(A) IS NOT APPROPRIATE.

A. Co-Pending Litigation Does Not Warrant Discretionary Denial (*Fintiv*)

Efficiency, fairness, and merits support institution despite the existence of co-pending litigation. *Apple v. Fintiv*, IPR2020-00019, Paper 11 (P.T.A.B. Mar. 20, 2020) (“*Fintiv*”).

1. Factor 1: Potential Stay

Patent Owner (“PO”) sued Petitioner for infringement of the ’058 patent and other patents approximately six months ago. *Jawbone Innovations, LLC v. Meta Platforms, Inc.*, No. 6:23-cv-00158, Dkt. No. 1 (W.D. Tex. Feb. 28, 2023). Meta answered Jawbone’s complaint, but nothing else has occurred in the litigation. *Id.*, Dkt. No. 13. There is no case schedule, and no scheduling conference is set, nor has any discovery been served.

Meta will soon move to transfer the litigation to Northern District of California. Two other Jawbone cases involving the same patents were previously transferred to NDCA on similar facts, and it is likely that Petitioner’s case will similarly be transferred. *In re Google LLC*, 58 F.4th 1379, 1383 (Fed. Cir. 2023); *Jawbone Innovations, LLC v. Amazon.com, Inc., et al.*, No. 2:21-cv-00435-JRG, Dkt. No. 60 (E.D. Tex. Sept. 21, 2022).

NDCA routinely grants motions to stay when IPRs have been instituted.

E.g., Uniloc USA Inc. v. LG Elecs., U.S.A. Inc., No. 18-cv-06737-JST, 2019 WL 1905161 (N.D. Cal. April 29, 2019); *Symantec Corp. v. Zscaler, Inc.*, No. 17-cv-04426-JST, 2018 WL 3539267 (N.D. Cal. July 23, 2018); *Twilio, Inc. v. TeleSign Corp.*, No. 16-cv-06925-LHK, 2018 WL 1609630 (N.D. Cal. April 3, 2018); *Robert Bosch Healthcare Sys., Inc. v. Cardiocom, LLC*, No. C-14-1575 EMC, 2014 WL 3107447 (N.D. Cal. July 3, 2014). Indeed, both prior Jawbone cases were stayed pending outcome of IPRs. No. 3:22-cv-06727-TLT, Dkt. No. 105; No. 3:23-cv-00466-TLT, Dkt. No. 137. While those IPRs are now terminated pursuant to settlements, institution decisions are imminent on new IPR petitions challenging each patent-at-issue in the litigation between Jawbone and Petitioner. *See infra* at Section VIII.B.

Because each petition relies upon the same grounds as a previously instituted proceeding that has since been terminated, institution and a litigation stay is highly likely. This factor weighs against discretionary denial.

2. Factor 2: Proximity of Trial to FWD

No trial date has been set. This factor weighs against discretionary denial. *See, e.g., Global Tel*Link Corp. v. HLFIP Holding, Inc.*, IPR2021-00444, Paper 14 at 16-19 (P.T.A.B. Jul. 22, 2021); *GAF Materials LLC v. Kirsch Research and Dev., LLC*, IPR2021-00192, Paper 14 at 12-13 (P.T.A.B. May 25, 2021); *Skechers*

U.S.A., Inc. v. Nike, Inc., IPR2021-00160, Paper 10 at 9-11 (P.T.A.B. May 19, 2021).

3. Factor 3: Investment in Parallel Proceeding

PO filed its complaint on February 28, 2023, and Petitioner filed its answer on May 30, 2023. The parties have invested little in the litigation. There is no case schedule. The parties have not conducted any discovery, exchanged proposed claim constructions, submitted claim construction briefs, nor served expert reports. Little will be invested in the litigation over the next several months. The “remaining investment” outweighs any investments made. *Samsung Elecs. Am. Inc. v. Snik LLC*, IPR2020-01428, Paper 10 at 11 (P.T.A.B. Mar. 9, 2021); *Huawei Techs. Co. v. WSOU Invs., LLC*, IPR2021-00226, Paper 10 at 13-14 (P.T.A.B. Jun. 10, 2021).

4. Factor 4: Overlapping Issues

If this IPR is instituted, Petitioner stipulates to not pursuing the petition’s grounds before the district court. This factor weighs against discretionary denial. *Sand Revolution II, LLC v. Cont’l Intermodal Grp.-Trucking LLC*, IPR2019-01393, Paper 24 at 11-12 (P.T.A.B. Jun. 16, 2020) (informative) (“Sand Revolution”); *Google LLC v. Jawbone Innovations, LLC*, IPR2022-00630, Paper 10 at 16-17 (P.T.A.B. Sept. 13, 2022).

5. Factor 5: The Parties

The parties are the same, but based on the facts above, it is unlikely that the parties' district court case will go to trial before this IPR reaches a final written decision. This factor is neutral. *See Google*, IPR2022-00630, Paper 10 at 14.

6. Factor 6: Other Circumstances

To Petitioner's knowledge, other than Google's petition with identical grounds (IPR2022-00623, filed March 8, 2022) and petitions from Amazon and Apple presenting different art and combinations (IPR2023-00222, filed Nov. 17, 2022; IPR2022-01021, filed May 16, 2022), which have all been terminated, the '058 patent has never been evaluated for anticipation or obviousness by the P.T.A.B, any court, or any jury. While two other IPR petitions challenging the '058 patent are pending institution (IPR2023-01110 and IPR2023-01119), those petitions rely upon different grounds than this petition. This petition's merits are particularly strong, which favors institution. *Fintiv*, IPR2020-00019, Paper 11 at 18.

Accordingly, *Fintiv* favors institution.

B. Existence of Previously Filed IPRs Does Not Warrant Discretionary Denial (*General Plastic*)

IPRs against the same claims of the '058 patent have been filed by Google (IPR2022-00623, filed Mar. 8, 2022), Apple (IPR2022-01021, filed May 16, 2022), Amazon (IPR2023-00222, filed Nov. 17, 2022), LG (IPR2023-01110, filed Jun. 23, 2023), and Sony (IPR2023-01119, filed Jun. 23, 2023). These petitions do

not warrant discretionary denial under *General Plastic Indus. Co., Ltd. v. Canon Kabushiki Kaisha*, IPR2016-01357, Paper 19 at 15-16 (P.T.A.B. Sept. 6, 2017) (precedential).

Meta and the prior petitioners are different parties and have no significant relationship. They are not co-defendants. They are competitors accused of infringement—in different actions pending in different courts—based on sales of different products. This weighs against denial. *NetNut Ltd. v. Bright Data Ltd.*, IPR2021-00465, Paper 11 at 9 (P.T.A.B. Aug. 12, 2021); *Mercedes-Benz USA, LLC v. Carucel Invs. L.P.*, IPR2019-01404, Paper 12 at 11-12 (P.T.A.B. Jan. 22, 2020); *Toshiba Am. Info. Sys., Inc. v. Walletex Microelecs. Ltd.*, IPR2018-01538, Paper 11 at 20 (P.T.A.B. Mar. 5, 2019).

The remaining factors consider timing issues, *e.g.*, when Petitioner became aware of prior art, whether the PO or Board responded to prior petitions, and the length of time that has passed. But the sequential IPR filings here were caused by PO's sequential district court filings and should not be held against Petitioner. Specifically, PO sued Samsung in May 2021 (2:21-cv-00186-JRG-RSP (E.D. Tex.)), waited four months before suing Google and Apple (6:21-cv-00985-ADA, 6:21-cv-00984-ADA (W.D. Tex.)), another two months to sue Amazon, and then waited another fourteen months, until Feb. 28, 2023, to sue Meta, Sony, and LG. The Board should not deny this petition merely because PO sued other defendants

earlier and independently. *Toshiba*, IPR2018-01538, Paper 11 at 21-22 (PO’s complaint about multiple IPRs “is not persuasive when the volume appears to be a direct result of its own litigation activity”); *id.* at 22.

Because PO’s litigation strategy—not Petitioner’s conduct—caused the filing of multiple IPRs, the Board should not deny institution under *General Plastic*.

VII. DISCRETIONARY DENIAL UNDER §325(D) IS NOT APPROPRIATE.

This petition should not be discretionarily denied under § 325(d). *Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6 (Feb. 13, 2020) (precedential). None of the prior art asserted in this petition was considered by the examiner during prosecution, and this petition’s arguments are not substantially the same as or cumulative of the arguments previously presented to the Office during original prosecution.

Though this petition presents identical grounds as Google’s petition (IPR2022-00623), Google’s proceeding terminated due to settlement before final written decision. Therefore, this petition should not be denied under 35 U.S.C. § 325(d). Furthermore, because this petition is substantively identical and includes the same grounds as the previously instituted Google IPR, Petitioner gains no

advantage and PO suffers no prejudice from arguments previously made by PO in IPR2022-00623.

VIII. Mandatory Notices

A. Real Parties-in-Interest Under 37 C.F.R. § 42.8(b)(1)

Meta Platforms, Inc. is the real party-in-interest.

B. Related Matters Under 37 C.F.R. § 42.8(b)(2)

To the best knowledge of Petitioner, the '058 patent is or has been involved in the following district court litigations and petitions for *inter partes* review:

Name	Number	Court	Filed
<i>Jawbone Innovations, LLC v. Amazon.com, Inc.</i>	2:22-cv-06727	N.D. Cal.	Nov. 29, 2021 (transferred from E.D. Tex. on Nov. 1, 2022, No. 2:21-cv-00435))
<i>Jawbone Innovations, LLC v. Apple Inc.</i>	6:21-cv-00984	W.D. Tex.	Sept. 23, 2021
<i>Jawbone Innovations, LLC v. Google LLC</i>	3:23-cv-00158	W.D. Tex.	Sept. 23, 2021 (transferred from E.D. Tex. on Feb. 1, 2023, No. 6:21-cv-00985)
<i>Jawbone Innovations, LLC v. Samsung Electronics Co., Ltd.</i>	2:21-cv-00186	E.D. Tex.	May 27, 2021
Petition for <i>Inter Partes</i> Review [Samsung & Apple]	IPR2022-01021	P.T.A.B.	May. 16, 2022
Petition for <i>Inter Partes</i> Review [Google]	IPR2022-00623	P.T.A.B.	Mar. 08, 2022

Name	Number	Court	Filed
Petition for <i>Inter Partes</i> Review [Amazon]	IPR2023-00222	P.T.A.B.	Nov. 17, 2022
<i>Jawbone Innovations, LLC v. Meta Platforms, Inc. d/b/a/ Meta</i>	6:23-cv-00158	W.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. ZTE Corp.</i>	2:23-cv-00082	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. Panasonic Holdings Corp.</i>	2:23-cv-00081	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. Guangdong Oppo Mobile Tele. Corp., Ltd. d/b/a OPPO</i>	2:23-cv-00079	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. LG Electronics, Inc.</i>	2:23-cv-00078	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. HTC Corp.</i>	2:23-cv-00077	E.D. Tex.	Feb. 28, 2023
<i>Jawbone Innovations, LLC v. Sony Elec., Inc.</i>	2:23-cv-01161	D.N.J.	Feb. 28, 2023
Petition for <i>Inter Partes</i> Review [Sony]	IPR2023-01119	P.T.A.B.	Jun. 23, 2023
Petition for <i>Inter Partes</i> Review [LG]	IPR2023-01110	P.T.A.B.	Jun. 23, 2023

In addition, institution decisions are imminent on new IPR petitions challenging each of the patents at issue in the district court litigation between Jawbone and Petitioner.

- IPR2023-01248, Paper 3 (P.T.A.B. July 26, 2023) ('543 patent, claims 1-26);
- IPR2023-01222, Paper 1 (P.T.A.B. July 20, 2023) ('543 patent, claims 1-26);

- IPR2023-01128, Paper 1 (P.T.A.B. June 27, 2023) (U.S. Pat. No. 8,326,611, claims 1-2, 5-28);
- IPR2023-01129, Paper 1 (P.T.A.B. June 27, 2023) (U.S. Pat. No. 8,326,611, claims 29, 31-44);
- IPR2023-01176, Paper 1 (P.T.A.B. July 6, 2023) (U.S. Pat. No. 8,326,611, claims 29-44);
- IPR2023-01177, Paper 1 (P.T.A.B. July 6, 2023) (U.S. Pat. No. 8,326,611, claims 1-28);
- IPR2023-01155, Paper 2 (P.T.A.B. July 6, 2023) (U.S. Pat. No. 8,326,611, claims 1-28);
- IPR2023-01156, Paper 1 (P.T.A.B. July 6, 2023) (U.S. Pat. No. 8,326,611, claims 29-44);
- IPR2023-01065, Paper 1 (P.T.A.B. June 15, 2023) (U.S. Pat. No. 10,779,080, claims 1-20);
- IPR2023-01110, Paper 2 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 7,246,058, claims 1-5);
- IPR2023-01119, Paper 1 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 7,246,058, claims 1-5);
- IPR2023-01130, Paper 1 (P.T.A.B. June 27, 2023) (U.S. Pat. No. 11,122,357, claims 1-20);
- IPR2023-01134, Paper 2 (P.T.A.B. June 29, 2023) (U.S. Pat. No. 11,122,357, claims 1-20);
- IPR2023-01153, Paper 1 (P.T.A.B. June 27, 2023) (U.S. Pat. No. 11,122,357, claims 1-20);
- IPR2023-01127, Paper 1 (P.T.A.B. June 27, 2023) (U.S. Pat. No. 8,019,091, claims 1-20);
- IPR2023-01154, Paper 2 (P.T.A.B. July 5, 2023) (U.S. Pat. No. 8,019,091, claims 1-20);

- IPR2023-01166, Paper 1 (P.T.A.B. July 5, 2023) (U.S. Pat. No. 8,019,091, claims 1-20);
- IPR2023-01083, Paper 1 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 8,321,213, claims 1, 3-13);
- IPR2023-01084, Paper 1 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 8,321,213, claims 14-36, 39-42);
- IPR2023-01111, Paper 2 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 8,321,213, claims 1-13);
- IPR2023-01112, Paper 2 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 8,321,213, claims 14-42);
- IPR2023-01117, Paper 1 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 8,321,213, claims 14-42);
- IPR2023-01118, Paper 1 (P.T.A.B. June 23, 2023) (U.S. Pat. No. 8,321,213, claims 1-13);
- IPR2023-01228, Paper 1 (P.T.A.B. July 20, 2023) (U.S. Pat. No. 8,503,691, claims 1-22, 29-40);
- IPR2023-01229, Paper 1 (P.T.A.B. July 24, 2023) (U.S. Pat. No. 8,503,691, claims 23-28, 41-46).

C. Lead and Back-Up Counsel Under 37 C.F.R. § 42.8(b)(3)

Meta provides the following designation of counsel:

Lead Counsel	Backup Counsel
Lisa Nguyen (Reg. No. 58,018) lisa.nguyen@allenoverly.com <u>Postal and Hand-Delivery Address:</u> Allen & Overy LLP 550 High Street	David Tennant (Reg. No. 48,362) david.tennant@allenoverly.com Alan Billharz (Reg. No. 79,532) alan.billharz@allenoverly.com Colby Davis (Reg. No. 81,575) colby.davis@allenoverly.com

Palo Alto, CA 94301 Telephone: 650-388-1724 Facsimile: 202-683-3999	<u>Postal and Hand-Delivery Address:</u> Allen & Overy LLP 1101 New York Ave. NW Washington, D.C. 20005 Telephone: (202) 683-3891 Facsimile: (202) 683-3999 Eric Lancaster (<i>pro hac vice</i> to be filed) eric.lancaster@allenoverly.com <u>Postal and Hand-Delivery Address:</u> Allen & Overy LLP 550 High Street Palo Alto, CA 94301 Telephone: 650-388-1724 Facsimile: 202-683-3999 Daniel Margolis (<i>pro hac vice</i> to be filed) daniel.margolis@allenoverly.com Grace Wang (Reg. No. 69,892) grace.wang@allenoverly.com <u>Postal and Hand-Delivery Address:</u> Allen & Overy, LLP 1221 6th Ave New York, NY 10020 Telephone: (212) 756-1143 Facsimile: (202) 683-3999
--	---

Pursuant to 37 C.F.R. § 42.10(b), a Power of Attorney accompanies this
petition.

D. Service Information Under 37 C.F.R. § 42.8(b)(4)

Service information for counsel is provided above. Meta consents to service by email at: AO_Meta_Jawbone@AllenOvery.com.

IX. Grounds for Standing

Meta certifies that the '058 patent is available for IPR; Meta is not barred or estopped from requesting IPR; and this petition is being filed within one year of service of the complaint in the co-pending district court litigation.

X. Conclusion

Petitioner requests institution of this IPR and cancellation of the claims.

Respectfully submitted,

August 31, 2023

/Lisa K. Nguyen/
Lisa K. Nguyen (Reg. No. 58,018)
lisa.nguyen@allenoverly.com
Allen & Overy LLP
550 High Street
Palo Alto, CA 94301
Telephone: 650.388.1724

Counsel for Petitioner,
META PLATFORMS, INC.

CERTIFICATE OF COMPLIANCE

Pursuant to 37 C.F.R. § 42.24(d), the undersigned certifies that foregoing
**PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-5 OF U.S.
PATENT NO. 7,246,058**, exclusive of the parts exempted as provided in 37
C.F.R. §42.24(a), contains 13,985 words and therefore complies with the type-
volume limitations of 37 C.F.R. §42.24(a).

Respectfully submitted,

August 31, 2023

/Lisa K. Nguyen/
Lisa K. Nguyen (Reg. No. 58,018)
lisa.nguyen@allenoverly.com
Allen & Overy LLP
550 High Street
Palo Alto, CA 94301
Telephone: 650.388.1724

Counsel for Petitioner,
META PLATFORMS, INC.

CERTIFICATE OF SERVICE

I hereby certify that a true and correct copy of the foregoing **PETITION FOR *INTER PARTES* REVIEW OF CLAIMS 1-5 OF U.S. PATENT NO. 7,246,058** and **EXHIBITS** are being served on September 1, 2023, via Federal Express overnight mail on counsel of record for U.S. Patent No. 7,246,058 as addressed below:

Mark Leonardo
Nutter McClennen & Fish LLP
Seaport West
155 Seaport Boulevard
Boston, MA 02210-2604

A courtesy copy is also being served on counsel for the patent holder in the pending litigation *Jawbone Innovations, LLC v. Meta Platforms, Inc.*, 6:23-cv-00158 (W.D. Tex., filed February 28, 2023):

Alfred R. Fabricant
Peter Lambrianakos
Vincent J. Rubino, III

FABRICANT LLP
411 Theodore Fremd Avenue,
Suite 206 South
Rye, New York 10580
Telephone: (212) 257-5797
Facsimile: (212) 257-5796
Email: ffabricant@fabricantllp.com
Email: plambrianakos@fabricantllp.com
Email: vrubino@fabricantllp.com
Email: eiturralde@fabricantllp.com
Email: jawbone@fabricantip.com

Raymond W. Mort, III
THE MORT LAW FIRM, PLLC
501 Congress Avenue, Suite 150
Austin, Texas 78701
Telephone: (512) 865-7950
Email: raymort@austinlaw.com

Email: rcowell@fabricantllp.com

August 31, 2023

Respectfully submitted,

/Lisa K. Nguyen/

Lisa K. Nguyen (Reg. No. 58,018)

lisa.nguyen@allenoverly.com

Allen & Overly LLP

550 High Street

Palo Alto, CA 94301

Telephone: 650.388.1724

Counsel for Petitioner,
META PLATFORMS, INC