

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

KRISP TECHNOLOGIES, INC.,
Petitioner,

v.

SANAS.AI INC.,
Patent Owner.

Case No. IPR2026-00275
Patent No. 11,715,457

PATENT OWNER'S PRELIMINARY RESPONSE

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	THE '457 PATENT	7
III.	CLAIM CONSTRUCTION.....	10
IV.	PETITIONER HAS NOT ESTABLISHED A PRIMA FACIE CASE OF OBVIOUSNESS BECAUSE IT IGNORES THE CLAIMED PIPELINE AND THE EXAMINER-IDENTIFIED INVENTIVE LIMITATION ..	11
	A. The Petition Does Not Demonstrate a Reasoned Motivation to Combine Prior Art References	17
	B. The Petition Does Not Demonstrate a Reasonable Expectation of Success	21
	C. The Petition Does Not Make a Particularized Showing Of Where Each Claim Element Is Disclosed In The Asserted Prior Art.....	23
	D. Institution Should Be Denied as an Inefficient Use of Office Resources	26
V.	PETITIONER’S CHALLENGES FAIL ON THE MERITS.....	27
	A. Ground 1 Fails on the Merits	28
	1. Ground 1 Independently Fails Because Prabhavalkar is Not Analogous Art.....	29
	2. Petitioner Fails to Show Context-Based Feature Extraction	31
	3. Petitioner Attempts to Patchwork the Examiner-Identified Inventive Concept	34
	4. Petitioner Fails to Show Speaker-Embedding Integration ..	35
	5. Petitioner Improperly Treats Fan’s Distinct Embodiments as Interchangeable	37
	6. Petitioner Fails to Show the Claimed Chunking Combination	40

7.	Ground 1 Dependent-Claim Challenges Do Not Cure the Independent-Claim Defects	43
B.	Grounds 2–7 Fail on the Merits	43
1.	Grounds 2–7 Fail Because They Rely on Ground 1.....	43
2.	Ground 2 Fails Because Accentron Does Not Cure the Missing Claimed Pipeline.....	44
3.	Ground 3 Fails Because Singh’s ASR Training Does Not Teach Accent-Reduced Linguistic Feature Extraction.....	47
4.	Ground 6 Fails Because Osuna Does Not Cure the Missing “Linguistic Features With a Reduced Accent” Limitation..	48
5.	Ground 5 Fails Because Bongiovi’s Loudness Processing Does Not Cure the Claimed Feature-Extraction and Output-Generation Defects	49
6.	Ground 4 Fails Because Wu’s VAD Does Not Cure the Missing Acoustic-Feature Extraction Architecture	50
7.	Ground 7 Fails Because Osuna and Wu Do Not Cure the Defects in Claim 14	50
VI.	CONCLUSION	52

TABLE OF AUTHORITIES

Page(s)

CASES

<i>ActiveVideo Networks, Inc. v. Verizon Commc'ns, Inc.</i> , 694 F.3d 1312 (Fed. Cir. 2012)	17
<i>Apple Inc. v. Ginko LLC</i> , IPR2025-01388, Paper 14 (March 2, 2026)	25
<i>In re Bigio</i> , 381 F.3d 1320 (Fed. Cir. 2004)	29
<i>Chevron Oronite Co. v. Infineum USA</i> , IPR2018-00923, Paper 9 (November 7, 2018)	26
<i>CyberSecure IPS, LLC v. Network Integrity Systems, Inc.</i> , IPR2025-01441, Paper 18 18 (May 26, 2026)	26
<i>Honeywell Int'l Inc. v. 3G Licensing, S.A.</i> , 124 F.4th 1345 (Fed. Cir. 2025)	21
<i>Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.</i> , 821 F.3d 1359 (Fed. Cir. 2016)	21
<i>In re Kahn</i> , 441 F.3d 977 (Fed. Cir. 2006)	17
<i>KSR Int'l Co. v. Teleflex Inc.</i> , 550 U.S. 398 (2007)	11, 17
<i>PAR Pharm., Inc. v. TWI Pharms., Inc.</i> , 773 F.3d 1186 (Fed. Cir. 2014)	21
<i>Xerox Corp. v. Bytemark, Inc.</i> , IPR2022-00624, Paper 9 (PTAB Aug. 24, 2022)	18

STATUTES

35 U.S.C. § 312(a)(3)	<i>passim</i>
35 U.S.C. § 314(a)	11

37 C.F.R. § 42.104(b)(4).....	23
37 C.F.R. § 42.107	1

PATENT OWNER’S EXHIBIT LIST

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2001	Complaint filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Jul. 7, 2025) (D.I. 1) (“Complaint”)
Ex2002	Answer filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Sep. 8, 2025) (D.I. 28) (“Answer and Counterclaims”)
Ex2003	Amended Complaint filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Sep. 22, 2025) (D.I. 35) (“Amended Complaint”)
Ex2004	Order on First Motion to Dismiss filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Sep. 29, 2025) (D.I. 37) (“Order on First Motion to Dismiss”)
Ex2005	Order Denying Krisp Motion to Dismiss First Amended Complaint and Granting in Part and Denying in Part Motion to Dismiss Defendant Krisp’s Counterclaims filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Dec. 1, 2025) (D.I. 58) (“Order Denying Krisp Motion to Dismiss First Amended Complaint”)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2006	Answer to Amended Complaint filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Dec. 15, 2025) (D.I. 60) (“Answer to Amended Complaint”)
Ex2007	Order Denying Motion to Dismiss filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Dec. 19, 2025) (D.I. 63) (“Order Denying Motion to Dismiss”)
Ex2008	Order Denying [Renewed] Motion for Judgment on Pleadings filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Feb. 23, 2026) (D.I. 72) (“Order Denying [Renewed] Motion for Judgment on Pleadings”)
Ex2009	Joint Case Management Statement filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Mar. 19, 2026) (D.I. 83) (“Joint Case Management Statement”)
Ex2010	Case Management Scheduling Order filed in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal. Mar. 26, 2026) (D.I. 92) (“Case Management Scheduling Order”)
Ex2011	Docket Sheet in <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> , No. 3:25-cv-05666 (N.D. Cal.) (“ <i>Sanas.AI Inc. v. Krisp Technologies, Inc.</i> Docket Sheet”)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2012	Coke Morgan Stewart, <i>Memorandum re Interim Processes for PTAB Workload Management</i> (Mar. 26, 2025), https://www.uspto.gov/sites/default/files/documents/InterimProcesses-PTABWorkloadMgmt-20250326.pdf . (“2025 Memo.”)
Ex2013	John A. Squires, <i>Memorandum re Additional Discretionary Institution Considerations—U.S. Manufacturing and Small Business Use of AIA Proceedings</i> (Mar. 11, 2026), https://www.uspto.gov/sites/default/files/documents/Additional_Discretionary_Institution_Considerations_US_Manufacturing_and_Small_Business_Use_of_AIA_Proceedings.pdf . (“2026 Memo.”)
Ex2014	<i>Magnolia Med. Techs., Inc. v. Kurin, Inc.</i> , No. IPR2026-00097, Paper 17 (PTAB May 14, 2026) (precedential) (“ <i>Magnolia</i> ”)
Ex2015	Scott R. Boalick, <i>Guidance on USPTO's Rescission of "Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation"</i> (Mar. 24, 2025)
Ex2016	[Intentionally Left Blank]
Ex2017	Sanas, LinkedIn, https://www.linkedin.com/company/sanastech/about/ (last visited May 13, 2026)
Ex2018	[Intentionally Left Blank]
Ex2019	[Intentionally Left Blank]

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2020	[Intentionally Left Blank]
Ex2021	[Intentionally Left Blank]
Ex2022	[Intentionally Left Blank]
Ex2023	[Intentionally Left Blank]
Ex2024	<i>About Sanas</i> , Sanas, https://www.sanas.ai/about-us (last visited May 16, 2026).
Ex2025	Krisp, LinkedIn, https://www.linkedin.com/company/krisphq (last visited May 13, 2026)
Ex2026	Krisp Team, <i>What's It Like to Work at Krisp?</i> (Sep. 20, 2021), https://krisp.ai/blog/working-at-krisp/ (last visited May 13, 2026) (“What’s it like to work at Krisp?”)
Ex2027	Arto Minasyan, <i>Krisp Voice AI Lab: Revolutionizing Voice Technology</i> , LinkedIn (posted Mar. 2026), https://www.linkedin.com/posts/artominasyan_eight-years-ago-we-started-with-a-simple-activity-7430219416912629760-8l6t (last visited May 13, 2026).
Ex2028	Krisp, <i>Introducing Accent Conversion for the Listener</i> , https://krisp.ai/blog/introducing-accent-conversion-for-the-listener/ (last visited May 20, 2026)

<u>EXHIBIT NO.</u>	<u>DESCRIPTION</u>
Ex2029	Arka Telecom, <i>YSU–Krisp AI Lab Opens at Yerevan State University</i> , https://arkatelecom.am/en/news/development/ysu_krisp_ai_lab_opens_at_yerevan_state_university/ (last visited May 20, 2026).
Ex2030	Careers at Krisp, Krisp, https://krisp.ai/careers/ (last visited May 14, 2026)
Ex2031	Krisp, Google Maps, https://www.google.com/maps (last visited May 13, 2026)
Ex2032*	U.S. Patent & Trademark Office, Enforcement and Non-Waiver of 37 C.F.R. § 42.104(b)(4) and Permissible Uses of General Knowledge in Inter Partes Reviews (July 31, 2025), https://www.uspto.gov/sites/default/files/documents/aap_a_memo_final_signed.pdf (“Enforcement and Non-Waiver Memo.”)

*** New Exhibit submitted herewith.**

I. INTRODUCTION

Pursuant to 37 C.F.R. § 42.107, Patent Owner Sanas.AI Inc. submits this Preliminary Response to the Petition for inter partes review of Claims 1–20 of U.S. Patent No. 11,715,457.

Institution here should be denied on the merits.

The '457 Patent claims an **end-to-end accent correction pipeline—chunking, dual context-based extraction, speaker embedding integration, and output generation**. The Examiner understood exactly that and allowed the claims because no prior art taught or suggested “extracting ... linguistic features with a reduced accent” from the input chunk and its context—a requirement central to the claimed architecture. Ex1001, Cl. 1(c).

Petitioner never identifies any reference—or combination—that discloses or renders obvious the claimed accent correction pipeline. Ex1002, at 115. Its underlying theory of prior art combinations is entirely illogical, cobbling together discrete fragments from three disconnected references directed to different problems and architectures, as illustrated by lynchpin Claim 1:

[1(Pre)] A method for real-time correction of an accent in a speech audio signal, the method being implemented by a computing device and comprising:

[1(a)] dividing the speech audio signal **into a stream of input chunks, an input chunk from the stream of input chunks**

including a predefined number of frames of the speech audio signal;

[1(b)] extracting, by an acoustic features extraction module from the input chunk and a context associated with the input chunk, acoustic features;

[1(c)] extracting, by a linguistic features extraction module from the input chunk and the context, linguistic features with a reduced accent

[1(d)] receiving a speaker embedding for a human speaker; and

[1(e)] generating an output chunk of an output audio signal based on the speaker embedding, the acoustic features, and the linguistic features.

Ex1001, Claim 1 (color coding supplied). The Petition fails to address—let alone map with particularity—the key limitation the Examiner identified as supporting allowance: extracting linguistic features with a reduced accent from both the input chunk and its context as part of the claimed pipeline. Ex1002, at 115.

The Petition sidesteps the Office’s reasoning and seeks to patch together three disparate references in an illogical manner:

(1) Feinauer (shown in black text) discloses *a context-driven accent and dialect modification framework* that operates using translation models, dictionaries, pairing databases, machine-learning models, and context selection to modify speech characteristics. Ex1004 at ¶¶[0047]–[0057], [0061]–[0078], [0104]–[0117]; the petition *does not assert that it discloses the claimed pipeline or, critically, any extraction of linguistic features with reduced accent from chunk-plus-context as required by the claims, let alone output any signal using these extracted features and speaker embedding;*

- (2) **Prabhavalkar (shown in blue bold text)**, by contrast, **is not directed to accent modification at all—it is an automatic speech recognition (“ASR”) system that processes speech frames through a sequence-to-sequence encoder-attender-decoder architecture to generate transcriptions**, and **does not generate accent-modified audio or any linguistic feature with a reduced accent**. Ex1005, Abstract, 1:15–25, 2:41–3:56, 6:56–7:12, 11:19–57. While **Prabhavalkar is being offered to show the use of chunking**, it is a totally different and irrelevant context;
- (3) Fan (shown in green underscore text), in turn, discloses voice-conversion architectures with selectable modes—such as style conversion, dialect conversion, and voice enhancement—each implemented through distinct networks, feature representations, and model pathways for synthesizing output speech. Ex1006 at ¶¶[0005]–[0007], [0045], [0055]–[0060], [0084]–[0095], [0016]–[0019]. But Fan does not disclose extracting linguistic features with a reduced accent from a chunk and its context in the manner required by the claims.

That color-coded patchwork of Feinauer, **Prabhavalkar**, and Fan does **not** constitute a coherent disclosure. At most, it identifies isolated elements drawn from distinct systems that operate on different inputs, employ different architectures, and yield different outputs. As such, it fails to establish any motivation to combine, reasonable expectation of success, or particularized mapping of claim elements.

This failure is particularly acute with respect to the limitation identified by the Examiner, which is nowhere disclosed in the prior art—individually or in any coherent combination.

The Petition does not explain **how** any of the identified elements would be integrated into a single system capable of performing the claimed sequence of operations; nor does it identify any teaching that aligns those components to generate an accent-corrected output audio signal based on acoustic features, linguistic features, and a speaker embedding. Instead, it reconstructs the claims from independent references without showing a system that actually performs the claimed method.

This pieced together combination is really an admission that **no single reference—or any coherent combination—teaches the claimed operation. Nor does any reference—or combination—teach the Examiner-identified requirement of extracting, from the input chunk and the context, “linguistic features with a reduced accent,” which remains the critical missing piece in the Petition’s theory.** Ex1002, at 115. Each reference addresses different data and architectures, and the Petition does not explain how those fundamentally different systems could be combined to perform the claimed real-time accent correction pipeline, which utilizes chunking, acoustic and linguistic feature extraction, and speaker embedding integration to generate an accent corrected audio signal.

Petitioner’s theory is essentially taking an accent and dialect modification framework/system (Feinauer), borrowing a **chunking mechanism** (**Prabhavalkar**), grafting a separate voice-conversion module (Fan), and then

declaring the result a real-time accent-correction engine. Those components may all fall under the broad umbrella of speech processing, but they operate on different data, pursue different objectives, and produce different outputs. Their superficial similarity does not mean the patchwork references fit together, let alone a coherent disclosure of the claimed method.

As the Petitioner does not identify any prior-art teaching of the claimed arrangement and operation of a real-time accent-correction pipeline—either in a single reference or a coherent combination—it does not cure the deficiency in the prior art identified by the Patent Office. Particularly, it leaves unaddressed the precise feature that the examiner cited in granting the claims: context-based extraction of linguistic features with a reduced accent integrated into the claimed real-time pipeline.

The remaining independent claim challenges similarly fail at least on these same defects. Claims 1, 14, and 20 recite the same operative invention across different classes—a method, a computing apparatus, and a computer-readable medium—and each requires the same integrated pipeline. *See* Ex1001 at Cl. 1, 14, 20. The Examiner allowed the claims because the prior art failed to teach that integrated chunking, extracting, and generating sequence, and the Petition offers no separate theory for Claims 14 and 20 beyond its defective Claim 1 mapping.

Ex1002 at 115. Because the same deficiencies infect the Petition's inadequate reading on all three independent claims, institution must be denied.

Grounds 2–7 only compound the flaws of Ground 1, and clearly do not cure them. They depend on the same deficient Feinauer-[Prabhavalkar-Fan](#) base combination and merely pile on even more disparate and irrelevant references that have no teachings to cure this defect, see Accentron (Ex1008), Singh (Ex1007), Wu (Ex1009), Bongiovi (Ex1010), and Osuna (Ex1011) purportedly for dependent-claim details. Putting aside whether they separately state a claim for the other grounds, for present purposes what matters is that none are offered to supply the missing Examiner-identified limitations or convert Petitioner's disconnected fragments into a coherent real-time accent-conversion pipeline.

This failure is dispositive: the Petition **does not** satisfy its burden under 35 U.S.C. § 312(a)(3) to show **with particularity where each claim element is found in the prior art**, instead reconstructing the claims from disconnected fragments.

II. THE '457 PATENT

The '457 Patent is directed to “real-time correction of an accent in a speech audio signal.” Ex1001 at Abstract. Claim 1 recites a specific pipeline for real-time accent correction which extracts acoustic features and linguistic features from chunks to generate an output based on these features and speaker embeddings. *Id.* at Cl. 1.

The claimed invention operates by continuously segmenting an incoming speech audio signal into time-based chunks and processing each chunk in real time through extraction. There are two extraction paths. The first extraction path focuses on capturing acoustic qualities of voice audio—such as pitch, energy, and whether speech is present. The second extraction path focuses on capturing linguistic content, identifying what is being said and minimizing the influence of the speaker’s accent. Importantly, each chunk is not analyzed in isolation; instead, the extraction paths rely on a running context of other portions of the audio, which helps to maintain continuity and improve the accuracy of the system’s understanding. The resulting acoustic and linguistic information is then combined with a speaker embedding, which represents the desired voice characteristics, and fed into a synthesis stage that generates a modified audio representation with a reduced accent. That representation is then converted back into audible speech,

producing output that preserves the original message while smoothing or removing accent-related features.

Claim 1's method requires:

[1(a)] dividing the speech audio signal **into a stream of input chunks, an input chunk from the stream of input chunks including a pre-defined number of frames of the speech audio signal;**

[1(b)] extracting, by an acoustic features extraction module from the input chunk and a context associated with the input chunk, acoustic features;

[1(c)] extracting, by a linguistic features extraction module from the input chunk and the context, linguistic features with a reduced accent;

[1(d)] receiving a speaker embedding for a human speaker; and

[1(e)] **generating an output chunk** of an output audio signal based on the speaker embedding, the acoustic features, and the linguistic features.

Ex1001 at Cl. 1.

As labeled in the Introduction, these elements define **a single, integrated real-time pipeline that must operate cohesively** by chunking, performing dual context-based extraction, integrating a speaker embedding, and generating outputs.

Critically, element **[1(c)]**—extracting “linguistic features with a reduced accent” from the input chunk and its context—was recognized by the Examiner as not taught or suggested by the prior art and as a basis for allowance. Ex1002, at

115.¹ This finding *reflects the Office’s determination that the prior art failed to disclose the very feature that differentiates the claimed real-time pipeline.*

Instead, the Petition simply ignores the context of the allowance-driving limitation while reconstructing the claims from disparate references—confirming that the Petition **improperly evades the Examiner’s reasoning and fails to meet its burden under 35 U.S.C. § 312(a)(3)** to show where this critical element is disclosed.

The Petition:

- **does not show** the asserted art teaches or suggests the integrated sequence of chunking, extracting, and generating an output chunk; and
- **does not sufficiently explain how** Prabhavalkar’s chunking or Fan’s style embedding vector could be incorporated into Feinauer’s accent and dialect modification framework/system;
- **improperly treats** Fan’s style separation and dialect conversion embodiments as interchangeable.

By failing to identify where the allowance-driving limitations are disclosed in the prior art, and by failing to explain how the proposed combination would

¹ The Office recognized that this novel pipeline applied to all the pending claims, including the other independent claims 14 and 20. Ex1002, at 115. In a similar manner, the Petition applies essentially identical analysis to all three independent claims. Pet. at 23-26, 41, 43.

perform those limitations as an integrated whole, the Petition lacks the particularized showing required by 35 U.S.C. § 312(a)(3) and should be rejected.

III. CLAIM CONSTRUCTION

Petitioner asserts that the claims should be interpreted under the *Phillips* standard and applies the plain and ordinary meaning of the claim terms as understood by a PHOSITA. Pet. at 3. For purposes of institution, Patent Owner agrees that no express claim construction is necessary because Petitioner's grounds fail under the plain and ordinary meaning of the claims. The institution determination on the merits turns on whether the prior art teaches "extracting, by a linguistic features extraction module from the input chunk and the context, linguistic features with a reduced accent", which the Petition does not come close to demonstrating. Accordingly, there is no claim construction dispute material to the institution decision. Patent Owner reserves the right to propose and address claim constructions should trial be instituted.

IV. PETITIONER HAS NOT ESTABLISHED A PRIMA FACIE CASE OF OBVIOUSNESS BECAUSE IT IGNORES THE CLAIMED PIPELINE AND THE EXAMINER-IDENTIFIED INVENTIVE LIMITATION

The Board may deny institution where a petitioner fails to satisfy its threshold burdens, which in the present case include demonstrating:

- (1) a reasoned motivation to combine prior art references,
- (2) a reasonable expectation of success, and
- (3) a particularized showing of where each claim element is disclosed in the asserted prior art.

35 U.S.C. §§ 103, 312(a)(3), 314(a); *see also KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007).

Each of these requirements is independently dispositive, and failure to satisfy any one of them warrants denial. Each independently show that there is not a reasonable likelihood of success that the petitioner would prevail. 35 U.S.C. § 314(a).

Here, the Petition fails because it never shows the claimed integrated real-time accent-correction pipeline. The reason the Petition cannot make this showing is that the alleged Feinauer–Prabhavalkar–Fan combination is inoperable. Petitioner does not identify a coherent prior-art architecture that performs the claimed method. Instead, it attempts to cobble together Feinauer’s high-level accent/dialect modification framework, Prabhavalkar’s ASR-specific chunking and temporal context, and Fan’s neural voice-conversion components. But no

amount of additional explanation can reconcile the architectural defects in that proposed combination.

For the chunking step **[1(a)]**, the proposed combination of **Prabhavalkar**'s ASR chunking with Feinauer's accent/dialect modification framework is not functional. Feinauer is directed to modifying accents and dialects in voice communications. Ex1004 ¶¶[0047]–[0057], [0061]–[0078], [0104]–[0117]. **Prabhavalkar**, by contrast, is an ASR reference. Its chunking is tied to a sequence-to-sequence speech-recognition model that processes frames and chunks to generate recognition outputs, including word elements and transcriptions—not modified audio. Ex1005, Abstract, 1:15–25, 2:41–3:56, 6:56–7:12, 11:19–57. Thus, **Prabhavalkar**'s chunking is not a deployable processing component for Feinauer's accent-modification architecture. The Petition does not sufficiently explain how Feinauer's word-level accent/dialect modifier would be restructured to receive and process **Prabhavalkar**-style ASR chunks, or why a PHOSITA would expect ASR chunking to operate within Feinauer's non-ASR modification framework.

For the acoustic-feature extraction step **[1(b)]**, the proposed combination also fails. Claim 1 requires extracting acoustic features from the input chunk and a context associated with that input chunk. Petitioner relies on Fan for the feature-extraction architecture and on **Prabhavalkar** for the chunk/context aspect. But

Prabhavalkar's "context" is ASR context used by an encoder-attender-decoder model to support speech recognition, not accent-corrected speech generation. Ex1005, 6:56–7:12, 9:8–10:65, 11:19–57. **Fan**'s feature-extraction models, in turn, operate within voice-conversion embodiments that extract PPGs, timbre, rhythm, accent, speaking speed, or style features for selected conversion modes. Ex1006 ¶¶[0007], [0085]–[0086], [0089]–[0095], [0126]–[0129]. Petitioner never explains how **Prabhavalkar**'s temporal ASR context would be incorporated into **Fan**'s feature-extraction architecture, or how that modified **Fan** component would then be grafted into Feinauer's accent/dialect modifier. Without that explanation, the Petition does not show a working acoustic-feature extraction module that extracts acoustic features from both the claimed input chunk and associated context.

For the linguistic-feature extraction step **[1(c)]**, the proposed combination is likewise deficient. Claim 1 requires extracting linguistic features with a reduced accent from the input chunk and the context. Petitioner points to **Fan**'s PPG-related disclosure for linguistic features and relies on **Prabhavalkar** for context. But **Fan**'s PPG extraction is not shown to operate on **Prabhavalkar**-style chunks plus prior-chunk context, and **Prabhavalkar**'s context is designed for ASR decoding rather than reduced-accent linguistic feature extraction. Ex1006 ¶¶[0007], [0085]–[0086]; Ex1005, 6:56–7:12, 18:1–41. Nor does Petitioner explain how **Fan**'s PPG features become "linguistic features with a reduced accent" in the claimed sense

when inserted into Feinauer’s framework. [Fan](#) describes PPGs as retaining content information by representing probabilities of phonemes, but the Petition does not show that [Fan](#)’s PPG extraction, as used in the alleged combination, removes or reduces accent information from the claimed chunk and context. Ex1006 ¶¶[0007], [0085]–[0086]. Thus, the proposed combination cannot coherently perform the claimed extraction of reduced-accent linguistic features from the input chunk and context.

For the speaker-embedding step **[1(d)]**, the proposed combination again fails. Petitioner relies on [Fan](#)’s style embedding vector as the claimed speaker embedding, while invoking Feinauer only as a high-level accent-modification backdrop. But [Fan](#)’s style embedding vector is not a free-floating value that can simply be inserted into Feinauer. It is generated through [Fan](#)’s own neural voice-conversion architecture, including a style separation model, sub-feature extraction models, user-selected weights, and a multihead attention structure. Ex1006 ¶¶[0087]–[0095]. Feinauer has no corresponding neural synthesis architecture for receiving that embedding, combining it with PPGs and acoustic features, and generating output speech. Ex1004 ¶¶[0047]–[0057], [0061]–[0078], [0104]–[0117]. Petitioner never explains how [Fan](#)’s embedding-generating neural network would interface with Feinauer’s accent/dialect modifier, which components would be retained or replaced, or what data representation would be passed between

them. The alleged combination therefore does not supply an operable speaker-embedding architecture.

The proposed combination also fails at the output-generation step **[1(e)]**. Claim 1 requires generating an output chunk of an output audio signal based on the speaker embedding, acoustic features, and linguistic features. Petitioner attempts to satisfy this limitation by combining **Prabhavalkar**'s chunk-based processing, **Fan**'s neural voice-conversion/vocoder architecture, and Feinauer's general accent/dialect modification framework. But those components do not fit together. **Prabhavalkar**'s chunk outputs are ASR recognition outputs, not output audio chunks. Ex1005, 6:4–28, 6:56–7:12, 18:1–41. **Fan** can generate speech through a neural voice-conversion pathway and vocoder, but that pathway depends on **Fan**'s own feature representations and network structure. Ex1006 ¶¶[0097], [0099], [0126]–[0129]. Feinauer, meanwhile, does not disclose a PPG-to-spectrum-to-vocoder pipeline; it modifies voice communications through its accent/dialect modification framework. Ex1004 ¶¶[0035]–[0047], [0053]–[0057], [0081]–[0086]. The Petition never provides an operable way to integrate **Prabhavalkar**'s chunks, **Fan**'s neural synthesis components, and Feinauer's accent/dialect modifier into a system that generates output audio chunks based on all three claimed inputs.

The further cobbling together of **Fan**'s distinct embodiments only compounds the problem. **Fan** discloses separate selectable modes, including style

conversion, dialect conversion, and voice enhancement, each with corresponding networks and pathways. Ex1006 ¶¶[0005]–[0007], [0016]–[0019], [0045], [0055]–[0060], [0084]–[0095]. Petitioner treats those embodiments as interchangeable, relying on Fan’s style-conversion components for embeddings and acoustic/style features while also invoking Fan’s dialect-conversion pathway for accent modification. But the Petition does not explain why a PHOSITA would combine those distinct Fan embodiments with each other, much less why that internally reconstructed Fan system would then be integrated into Feinauer and **Prabhavalkar**. This is not routine reprogramming; it is an unsupported wholesale reconstruction of the ’457 Patent’s claimed architecture using isolated fragments from unrelated systems. And nowhere does the Petition explain how a PHOSITA would—unguided by hindsight—combine the disparate approaches of the three references to construct the particular integrated real-time accent reduction pipeline claimed in the ’457 Patent.

Accordingly, the alleged Feinauer–**Prabhavalkar**–Fan combination does not teach or render obvious the claimed integrated pipeline. Feinauer supplies only a high-level accent/dialect modification framework, **Prabhavalkar** supplies ASR-specific chunking and context, and Fan supplies neural voice-conversion components from distinct embodiments. The Petition does not provide an operable path for those components to work together to perform the claimed sequence of

steps. Because Petitioner fails to show a coherent architecture, a reasoned motivation to combine, or a reasonable expectation of success, Ground 1 fails.

Each of these incompatibilities separately renders the proposed combination inoperable as the claimed pipeline. The result is not an integrated system that chunks, extracts, and generates as the claims require but a set of disconnected components that cannot be assembled into the claimed invention without rebuilding the system. These failures are carried through to all the grounds which rely upon this defective analysis.

A. The Petition Does Not Demonstrate a Reasoned Motivation to Combine Prior Art References

First, obviousness *requires more than identifying references* that could be combined; a petitioner *must articulate a reasoned motivation to combine those references, supported by a rational underpinning*. *KSR Int’l Co. v. Teleflex Inc.*, 550 U.S. 398, 418 (2007); *In re Kahn*, 441 F.3d 977, 988 (Fed. Cir. 2006).

Conclusory assertions—such as invocations of “known techniques”, “reprogramming”, or “predictable results”—are insufficient where, as here (*see e.g.* Pet. at 17, 20, 23, 25, 31, 37, 39, 48, 51, 55, 60, 63, 67-68, 71, 74), the Petition does not explain how the specific disclosures would be integrated into a working combination. *See ActiveVideo Networks, Inc. v. Verizon Commc’ns, Inc.*, 694 F.3d 1312, 1327–28 (Fed. Cir. 2012) (rejecting generic expert testimony that “failed to

explain how specific references would be combined, which combination(s) of elements in specific references would yield a predictable result, or how any specific combination would operate or read on the asserted claims”).

Petitioner’s declarant Mr. Schmandt (Ex1003) should be accorded little weight since he merely repeats unsupported assertions without providing additional supporting analysis or evidence. *Xerox Corp. v. Bytemark, Inc.*, IPR2022-00624, Paper 9 at 15–17 (PTAB Aug. 24, 2022) (precedential) (holding that the expert testimony should be “entitled to little weight” because the declaration “is conclusory and unsupported” and “adds little to the conclusory assertion for which it is offered to support”).

Here, as discussed in Section V(A), Petitioner does not provide a reasoned explanation for how Feinauer, **Prabhavalkar**, and Fan would be combined into a coherent system that performs the claimed real-time accent-correction pipeline. While the Petition identifies purported reasons why a PHOSITA would have combined Feinauer, **Prabhavalkar**, and Fan, those assertions are conclusory and fail to address the technical incompatibilities noted above that such a combination entails.

With respect to **Prabhavalkar**, the Petition argues that Feinauer’s high-level disclosure of processing speech “portions” could be implemented using

Prabhavalkar's ASR chunking, supposedly through routine "reprogramming." Pet. at 17. But the Petition never explains how **Prabhavalkar**'s chunking and temporal-context framework could function within Feinauer's accent/dialect modification architecture. That incompatibility is fatal. Feinauer modifies voice communications using an accent and dialect modifier, translation models, dictionaries, pairing databases, and context-based selection. Ex1004 ¶¶[0047]–[0057], [0061]–[0078], [0104]–[0117]. **Prabhavalkar**, by contrast, uses chunks in an ASR system that processes speech frames through an encoder-attender-decoder architecture to generate recognition outputs, such as word elements or transcriptions—not modified speech audio. Ex1005, Abstract, 1:15–25, 6:56–7:12, 11:19–57, 18:1–41. The Petition offers only the unsupported assertion that **Prabhavalkar**'s ASR chunking can be "readily" incorporated into Feinauer's non-ASR system, without sufficient detail, logic, or evidentiary support to make the necessary bridge to justify institution.

The same defect appears in the Petition's reliance on Fan. Fan supplies the specific neural voice-conversion components that Petitioner needs for feature extraction, embeddings, and output generation. But Fan's architecture is not a plug-in component for Feinauer. Fan extracts PPGs and style features within distinct voice-conversion embodiments, including style conversion and dialect conversion, each with its own model pathway. Ex1006 ¶¶[0007], [0085]–[0086],

[0089]–[0095], [0126]–[0127]. The Petition treats those distinct Fan embodiments as interchangeable, asserting without support or analysis that Fan’s style-separation and embedding components would be equally applicable to Fan’s dialect-conversion pathway and then further insertable into Feinauer. Pet. at 18–31. But the Petition never explains how Fan’s neural feature representations, style embedding vector, or synthesis pathway could be inserted into Feinauer’s word-level accent/dialect modification framework. Thus, the Petition fails to sufficiently explain not only that the proposed combination would make the claimed result, but also why a PHOSITA would have been motivated to make such a combination in the first instance.

The Petition’s treatment of Fan’s speaker embedding highlights the problem. Feinauer provides, at most, a high-level backdrop for modifying speech characteristics, while Petitioner’s actual speaker-embedding architecture is imported from Fan. Fan’s style embedding is generated through Fan’s own neural architecture, including style separation, sub-feature extraction, and multihead attention. Ex1006 ¶¶[0089]–[0095]. Feinauer, however, does not disclose a corresponding neural synthesis pipeline capable of receiving that embedding, combining it with PPGs and acoustic features, and producing output audio. Ex1004 ¶¶[0047]–[0057], [0061]–[0078]. The Petition never explains what component in Feinauer would receive Fan’s embedding, what representation would be passed

between the systems, or how [Fan](#)'s embedding-generating model would interface with Feinauer's accent/dialect modifier.

Because the Petition relies on conclusory assertions of convenience and ease of modification rather than reasoned analysis addressing the incompatibility, it fails to establish the required motivation to combine.

B. The Petition Does Not Demonstrate a Reasonable Expectation of Success

Second, a petitioner must demonstrate *a reasonable expectation of success for the proposed combination, not merely a motivation to attempt it*. *Intelligent Bio-Sys., Inc. v. Illumina Cambridge Ltd.*, 821 F.3d 1359, 1367–68 (Fed. Cir. 2016); *PAR Pharm., Inc. v. TWI Pharms., Inc.*, 773 F.3d 1186, 1193 (Fed. Cir. 2014). While a combination need not be “the preferred, or the most desirable, combination,” the petitioner must show “there is something in the prior art as a whole to suggest the desirability, and thus the obviousness, of the claimed invention.” *Honeywell Int'l Inc. v. 3G Licensing, S.A.*, 124 F.4th 1345, 1355 (Fed. Cir. 2025) (quotation omitted).

Here, as discussed in Section V, the Petition gives a PHOSITA no adequate basis to expect that the proposed combination would succeed in producing the integrated real-time accent-correction pipeline that Claims 1, 14, and 20 require.

Nor does the Petition explain how these fundamentally different architectures could be integrated to perform the claimed operations, particularly the Examiner-identified limitation that was found missing from the prior art.

Feinauer is a context-driven accent and dialect modification system that uses an accent and dialect modifier, translation models, dictionaries, pairing databases, and context selection to modify speech characteristics. **Prabhavalkar** is an ASR system that processes speech frames through a sequence-to-sequence encoder-attender-decoder architecture to generate recognition outputs, such as word elements or transcriptions, not modified speech audio. **Fan** is a neural voice-conversion system with distinct selectable modes—such as style conversion and dialect conversion—implemented through separate model pathways, including PPG extraction, style separation, embedding generation, mel-feature generation, and vocoder-based output. See Ex1004 ¶¶[0047]–[0057], [0061]–[0078], [0104]–[0117]; **Ex1005, Abstract, 1:15–25, 6:56–7:12, 11:19–57**; Ex1006 ¶¶[0005]–[0009], [0084]–[0095], [0126]–[0127].

These are not merely different references but materially different architectures, and the Petition never explains why a skilled artisan would have expected Feinauer’s word-level accent/dialect modification framework, **Prabhavalkar**’s ASR-specific chunking and attention-based context, and **Fan**’s neural PPG/style-embedding/synthesis pipeline to interoperate within a single end-

to-end real-time accent-correction system. An expectation of success asserted at the broad level of generality of the Petition, untethered from any substantive analysis bridging these architectural differences, is no expectation of success at all, and cannot sustain the obviousness ground.

C. The Petition Does Not Make a Particularized Showing Of Where Each Claim Element Is Disclosed In The Asserted Prior Art

Third, the petition must identify “**with particularity**” **where** each claim element is disclosed in the prior art and **explain how** the asserted grounds render the claims unpatentable. 35 U.S.C. § 312(a)(3); 37 C.F.R. § 42.104(b)(4). The Office has made clear that missing claim limitations cannot be supplied by general knowledge, common sense, or conclusory expert testimony. Ex2032 at 1.

Here, the Petition maps the individual elements of independent Claims 1, 14, and 20 to scattered disclosures across three references—Feinauer for the general accent/dialect modification framework, [Prabhavalkar](#) for chunking and temporal context, and [Fan](#) for feature extraction, speaker embeddings, and output generation—but never makes a particularized showing that those pieces combine into the claimed pipeline as an operable whole. Pet. at 16–45. That failure is especially clear for elements **[1(b)]** and **[1(c)]**. Those limitations do not merely require extracting some acoustic or linguistic information from speech; they require extracting acoustic features and linguistic features with a reduced accent

from the input chunk and the context associated with that chunk, **as part of a real-time accent-correction pipeline.**

Neither **Prabhavalkar** nor Fan teaches that integrated limitation.

Prabhavalkar provides ASR-specific chunking and context for a recognition system that generates word elements or transcriptions, not modified speech audio. Ex1005, Abstract, 6:56–7:12, 11:19–57, 18:1–41. Fan, by contrast, extracts PPGs and style/acoustic features within neural voice-conversion embodiments, but the Petition does not show that Fan extracts those features from **Prabhavalkar**-style input chunks together with prior-chunk context, nor that Fan's PPGs constitute linguistic features with a reduced accent in the claimed context. Ex1006 ¶¶[0007], [0085]–[0086], [0089]–[0095], [0126]–[0127]. Thus, pointing to **Prabhavalkar** for chunking/context and Fan for PPG or style extraction is not sufficient. The Petition needed to explain **how** **Prabhavalkar**'s ASR temporal-context mechanism is **integrated** into Fan's feature-extraction architecture and then **incorporated** into Feinauer's accent/dialect modification framework. **The Petition provides no such explanation.**

That omission is fatal under § 312(a)(3). The claims require the chunking, context-based acoustic-feature extraction, context-based reduced-accent linguistic-feature extraction, speaker embedding, and output-audio-chunk generation steps to operate **within a single integrated system.** The Examiner allowed the claims

because the claimed integrated sequence was missing from the prior art. The Petition never supplies the missing integration. It instead identifies isolated parts in separate references and assumes the claimed system follows.

Identifying where each limitation appears in some reference, standing alone, does not discharge Petitioner's burden. The claim requires that the chunking, extraction, speaker-embedding, and output-generation steps operate together in a real-time accent-correction pipeline, and it is that integration—the very feature the Examiner identified as missing from the prior art—for which the Petition offers no particularized analysis.

Thus, Petitioner cobbles the limitation together from separate and disparate sources **without showing**, element by element, how Feinauer's accent/dialect modification framework, [Prabhavalkar](#)'s ASR chunking and temporal context, and [Fan](#)'s neural feature-extraction and voice-conversion components combine to form a coherent pipeline that performs the claimed sequence of operations, let alone that such a combination would result in the claimed pipeline.

Apple Inc. v. Ginko LLC underscores that, where, as here, a petition fails to map claim elements with particularity, denial at institution is warranted. IPR2025-01388, Paper 14 at 3. There, the Director emphasized that a petitioner must, “from the onset,” provide an understandable, element-by-element explanation grounded in the prior art. Petitioner's showing here falls short for the same reasons: it neither

identifies a prior-art disclosure of the claimed **chunking, context-based extraction of reduced-accent linguistic features, and output-audio-chunk generation within an integrated system**, nor explains how the proposed combination performs **that requirement within the claimed pipeline**.

D. Institution Should Be Denied as an Inefficient Use of Office Resources

Finally, denial is warranted where institution would not serve the efficient administration of the Office. The Director has confirmed that the Board may deny institution where only a subset of claims or grounds has potential merit and proceeding to trial would waste resources. *CyberSecure IPS, LLC v. Network Integrity Systems, Inc.*, IPR2025-01441, Paper 18 18 (May 26, 2026) at 2–3 (citing *Chevron Oronite Co. v. Infineum USA*, IPR2018-00923, Paper 9 at 10–11 (November 7, 2018) (informative)).

Here, all asserted grounds depend on the same defective Feinauer, **Prabhavalkar**, and **Fan** combination, and the additional references in Grounds 2–7 do not cure the deficiencies in the independent claims.

Instituting trial would therefore require the Board to expend resources on multiple grounds that fail to address the dispositive issue.

V. PETITIONER’S CHALLENGES FAIL ON THE MERITS

Petitioner asserts seven obviousness grounds:

<u>Ground</u>	<u>Claims Challenged</u>	<u>References Relied Upon</u>	<u>Petition Citation</u>
Ground 1	1-4, 8, 14-17, 20	Feinauer, Prabhavalkar , Fan	Pet. 13
Ground 2	9–11	Feinauer, Prabhavalkar , Fan , Accentron	Pet. 44
Ground 3	6, 19	Feinauer, Prabhavalkar , Fan , Singh	Pet. 57
Ground 4	7	Feinauer, Prabhavalkar , Fan , Wu	Pet. 61
Ground 5	12–13	Feinauer, Prabhavalkar , Fan , Bongiovi	Pet. 64
Ground 6	5	Feinauer, Prabhavalkar , Fan , Osuna	Pet. 72
Ground 7	18	Feinauer, Prabhavalkar , Fan , Wu, Osuna	Pet. 74

This Preliminary Response focuses on Ground 1 because Ground 1 supplies the base combination that every other asserted ground, and every claim, requires. Ground 1 asserts that claims 1, 3–5, 8–10, 14, 16–17, and 20 are obvious over Feinauer, [Prabhavalkar](#), and [Fan](#). Pet. at 13. Grounds 2-7 merely add Accentron, Singh, Wu, Bongiovi, or Osuna for dependent-claim limitations, while retaining

the same Feinauer-**Prabhavalkar**-**Fan** combination for the independent-claim limitations. Pet. 44, 57, 61, 64, 72, 74. Thus, if Ground 1 fails to show the claimed alignment in independent claims 1, 14, and 20, Grounds 2-7 fail as well.

A. Ground 1 Fails on the Merits

Ground 1 fails because Petitioner attempts to reconstruct the '457 Patent through hindsight rather than identify a teaching or reasoned combination that a PHOSITA at the time of the invention would have recognized in the prior art. The Petition relies on three references that perform meaningfully different functions using materially distinct architectures. Feinauer modifies speech using an accent and dialect modifier system. Ex1004 at ¶[0047]. **Prabhavalkar** recognizes speech using ASR architecture that outputs transcriptions. Ex1005 at Abstract, 1:15–25, 2:41–3:56. **Fan** converts speech based on the selection of a particular voice-conversion mode. Ex1006 at ¶[0005].

The Petition never bridges those differences. It does not identify the combined architecture, explain which components are retained or discarded, or show how outputs from one system become usable inputs to another. In particular, the Petition does not sufficiently explain how Feinauer's accent modifier model would consume **Prabhavalkar**'s ASR chunks, how **Fan**'s PPG extractor would receive **Prabhavalkar**'s previous-chunk context, or how **Fan**'s embedding-based

neural conversion pathway would be integrated into Feinauer’s modifier. Those deficiencies defeat both motivation to combine and reasonable expectation of success.

Accordingly, the challenges to independent claims 1, 14, and 20 fail on the merits.

1. Ground 1 Independently Fails Because Prabhavalkar is Not Analogous Art

Petitioner’s Ground 1 fails at the threshold because **Prabhavalkar** (Ex1005) is not analogous art. A reference is analogous only if it is from the same field of endeavor or is reasonably pertinent to the problem faced by the inventor. *In re Bigio*, 381 F.3d 1320, 1325 (Fed. Cir. 2004) (“Two separate tests define the scope of analogous prior art: (1) whether the art is from the same field of endeavor, regardless of the problem addressed and, (2) if the reference is not within the field of the inventor’s endeavor, whether the reference still is reasonably pertinent to the particular problem with which the inventor is involved”). **Prabhavalkar** satisfies neither requirement.

Prabhavalkar is not in the same field of endeavor as the ’457 Patent. The ’457 Patent is directed to real-time accent correction and audio generation through an integrated pipeline that divides speech into input chunks, extracts acoustic features and reduced-accent linguistic features from the input chunk and context, receives a speaker embedding, and generates an output audio chunk based on those

inputs. Ex1001, Abstract, Cl. 1. **Prabhavalkar**, by contrast, is an automatic speech recognition (“ASR”) reference that processes speech frames through a sequence-to-sequence encoder-attender-decoder architecture to generate recognition outputs, such as word elements or transcriptions. Ex1005, Abstract, 1:15–25, 2:41–3:56, 6:56–7:12, 11:19–57. Its chunking and context are used to improve streaming recognition, not to modify accents, extract reduced-accent linguistic features, receive speaker embeddings, synthesize accent-corrected audio, or generate output audio chunks.

Prabhavalkar is also not reasonably pertinent to the problem addressed by the ’457 Patent. The ’457 Patent uses chunks and context as part of a real-time accent-correction pipeline that extracts acoustic and linguistic information and uses those features, together with a speaker embedding, to generate accent-corrected output audio. Ex1001, Abstract, Cl. 1. **Prabhavalkar** addresses a different problem: enabling an ASR model to process current and prior speech frames without waiting for a complete utterance and to output transcription-related recognition results. Ex1005, 6:56–7:12, 9:8–10:65, 11:19–57, 18:1–41. A PHOSITA seeking to build the claimed accent-correction pipeline would not have looked to **Prabhavalkar**’s ASR chunking and temporal context to solve the problem of extracting reduced-accent linguistic features for downstream audio generation.

Because [Prabhavalkar](#) is not analogous art, it cannot supply the claimed chunking or context limitations in Ground 1. Petitioner’s reliance on [Prabhavalkar](#) is hindsight-driven, and Ground 1 fails at the threshold.

2. Petitioner Fails to Show Context-Based Feature Extraction

Claim 1 requires two context-based extraction steps. The acoustic features extraction process must extract acoustic features “from the input chunk and a context associated with the input chunk.” Ex1001 at Cls. 1, 14, 20. The linguistic features extraction module must extract linguistic features with a reduced accent “from the input chunk and the context.” *Id.* Petitioner cannot satisfy those limitations through manipulation of chunking in [Prabhavalkar](#) and feature extraction in [Fan](#). Thus, the proposed combination **would not result in the claimed invention.**

This failure is most acute for element **[1(c)]**, which was specifically identified by the Examiner as missing from the prior art: extracting linguistic features with a reduced accent from both the input chunk and the associated context. Petitioner’s reliance on [Prabhavalkar](#)’s ASR attention and [Fan](#)’s voice-conversion feature extraction does not supply that requirement.

As discussed above, [Prabhavalkar](#)’s context is ASR context. Ex1005 at 9:20–36. Its encoder generates encoded feature representations; its attender generates a context vector by attending to encoder outputs; and its decoder uses

that context vector to output probability distributions over word elements for transcription. *Id.* at 9:8–10:65, 11:19–57. In streaming implementations, **Prabhavalkar**’s attender may compute attention based on speech frames in the current chunk and one or more previous chunks, including at least ten previous chunks. *Id.* at 6:56–7:12. That context supports recognition, not accent-corrected audio generation.

Fan’s feature extraction operates differently. **Fan**’s voice feature extraction model extracts PPG features from a first voice. Ex1006 at ¶¶[0007], [0067], [0084]–[0086]. But **Fan** does not disclose extracting PPGs or style/acoustic features from an input chunk plus a context. The attention mechanism disclosed in **Fan**, unlike that of **Prabhavalkar**, operates without a temporal dimension. *Id.* at ¶[0094]. **Fan**’s system focuses on weighting an extracted set of feature vectors (*e.g.*, timbre vector, prosody vector, accent vector) rather than analyzing a temporal sequence of audio chunks. *Id.* at ¶[0090].

Petitioner’s reliance on “attention” does not bridge the gap. Pet. 18-26. **Prabhavalkar**’s attention is temporal ASR attention. Ex1005 at 7:5-16. It accepts as input speech frames in the current chunk or previous chunks, and outputs at least one probability distribution of word elements for each chunk. *Id.* at 6:56–7:12. **Fan**’s multihead attention accepts as input a feature vector (*e.g.* timbre feature vector) and weighted sub-feature vectors, and aligns them to output a style

embedding vector. Ex1006 at ¶[0094]. These mechanisms serve different purposes, use different inputs, and generate different outputs. *See* Ex1005 at 2:44-47, 6:56-7:16; Ex1006 at ¶¶[0012], [0094]. The Petition never explains how **Prabhavalkar**'s temporal ASR lookback window would be inserted into Fan's PPG extraction or style separation models. And unlike **Prabhavalkar**, Fan does not require or utilize a lookback structure; instead, it extracts all necessary information from a single voice sample in a unified operation. *Id.* at ¶¶[0005]–[0008], [0066]–[0071]; *see also id.* at ¶¶[0016], [0020], [0084], [0087]. As Fan's attention mechanism acts upon an already computed set of features, it functions independent of the prior audio history.

Given these fundamental differences, the Petition fails to provide a persuasive rationale for why a PHOSITA would seek to integrate **Prabhavalkar**'s temporal chunking into Fan's weighting architecture. *See* Pet. at 22-23, 25-26. The two systems employ “attention” in fundamentally different capacities: **Prabhavalkar** uses it to navigate an evolving temporal sequence, whereas Fan uses it to weight static acoustic embeddings. Because these mechanisms serve distinct functional roles and operate on different types of data, the mere commonality of the term “attention” is insufficient to motivate the combination of references.

3. Petitioner Attempts to Patchwork the Examiner-Identified Inventive Concept

A central problem with Ground 1 is that it attempts to reconstruct the very feature the Examiner recognized as missing from the prior art. During prosecution, the Examiner identified “extracting, by a linguistic features extraction module from the input chunk and the context, linguistic features with a reduced accent” as not taught, disclosed, or suggested by the prior art of record in combination with the other limitations. Ex1002 at 115.

That limitation is not an isolated drafting detail. It is part of the claimed real-time accent-correction architecture, which requires chunked processing, context-based extraction of both acoustic and linguistic features, linguistic features with a reduced accent, receipt of a speaker embedding, and generation of an output audio chunk based on that claimed combination of inputs. Ex1001 at Cl. 1. Petitioner does not identify any independent reference containing that architecture, nor does Petitioner provide a reasoned rationale explaining how a PHOSITA would have combined the prior art to achieve it, revealing the Feinauer-**Prabhavalkar**-Fan **would not result in the claimed invention.**

Instead, Ground 1 pulls disconnected pieces from different systems: Feinauer for a general accent and dialect modification system, **Prabhavalkar** for chunking and temporal context, and Fan for style/acoustic features, extraction, and embeddings. See Pet. at 13-34. None of these references teaches the independent

claims, particularly the Examiner-identified limitation, and the Petition does not sufficiently explain how the combination would achieve this result.

The Petition offers impermissible hindsight reconstruction. The Examiner already recognized that the prior art did not teach the claimed extraction of linguistic features with a reduced accent from both the input chunk and context. Petitioner's Ground 1 repeats that same failure.

4. Petitioner Fails to Show Speaker-Embedding Integration

Claim 1 requires "receiving a speaker embedding for a human speaker." Ex1001 at Cl. 1. Petitioner relies on [Fan](#)'s style embedding vector. Pet. at 26-31. But [Fan](#)'s style embedding is not a generic value that can be inserted into any accent-modification system. As the style embedding vector is not pre-computed, or capable of being freely extracted and implemented in other systems, the combination proposed by the Petition would require incorporating the complete style separation model. However, the Petition provides no explanation as to whether Feinauer's system and [Fan](#)'s style separation model are compatible and if they are, how the incorporation would be effectuated. Accordingly, this **would not result in the claimed invention.**

[Fan](#)'s style embedding vector is the output of a specific neural-network process. In the style conversion embodiment, the style separation model extracts timbre and sub-feature vectors from a third voice; user-adjusted weights may be

applied to sub-features such as prosody, accent, and speaking speed; and a multihead attention structure outputs a style vector, for example a 256-dimensional style embedding vector. Ex1006 at ¶¶[0087]–[0095]. That embedding is then used with a PPG feature in [Fan](#)’s voice fusion model. Ex1004 ¶¶[0008]–[0009], [0094]–[0095].

Feinauer is organized differently. Its accent and dialect modifier relies on word-level mapping and static pre-trained parameters. Ex1004 at ¶¶[0047]–[0057], [0061]–[0078], [0104]–[0117].

The Petition never explains how [Fan](#)’s embedding-producing neural network interfaces with Feinauer’s modifier because there is no indication that Feinauer has the infrastructure to process high-dimensional continuous vectors. Incorporating [Fan](#)’s embeddings would require fundamental architectural components that are entirely absent from Feinauer’s disclosure.

In proposing the combination of Feinauer and [Fan](#), the Petition neglects to confirm whether Feinauer has the required hardware to run [Fan](#)’s software. Instead, the Petition simply asserts that “[a] PHOSITA would have recognized the modification constitutes applying a known technique (Feinauer and [Fan](#)’s steps of acoustic feature extraction) to a known device ([Fan](#)’s hardware and software modules) with predictable results”. Pet. at 20. There is nothing included that

remotely suggests that Feinauer’s hardware is compatible with Fan’s software. The provided motivation to combine is lacking.

Nor does the Petition show the required downstream use of the alleged speaker embedding in element **[1(e)]**. Claim 1 does not merely require the existence of some embedding in isolation; it requires generating an output chunk based on the speaker embedding, the acoustic features, and the linguistic features. Because Petitioner never establishes the claimed **[1(b)]** and **[1(c)]** feature extractions—especially the Examiner-identified linguistic features with a reduced accent—it also cannot show that Fan’s alleged embedding is used in the claimed output-generation combination.

Hence, the Petitioner fails to establish the speaker-embedding limitation in the proposed combination. The Petition only explains that it would have been obvious to a PHOSITA to combine the “method of Feinauer ... [with a] speaker embedding (style embedding vector) to modify the accent (dialect) of the output audio signal.” Pet. at 30. There is no discussion of the means by which the model would be incorporated, and this is insufficient to establish an effective combination of references.

5. Petitioner Improperly Treats Fan’s Distinct Embodiments as Interchangeable

Fan describes distinct selectable modes with corresponding voice conversion networks: style conversion, dialect conversion, and voice enhancement. Ex1006 at

¶¶[0005]–[0006], [0045], [0055]–[0060]. The Petition treats those embodiments as a single pool of interchangeable components. Pet. at 18-23. Fan does not (Ex1006 at ¶¶[0009], [0018]), and as a result, **the claimed invention does not follow.**

For Claim **[1(b)]**, the Petition relies upon Fan’s style separation model to meet the limitation requiring the presence of an “acoustic features extraction module.” Pet. at 18-23. However, the style separation model operates in completely different ways in Fan’s style enhancement embodiment and dialect conversion embodiment. See Ex1006 at ¶¶[0009], [0018]. To overcome this hurdle, the Petition simply states, without further explanation, that “[a] PHOSITA would have recognized descriptions related to any hardware and/or software related to Fan’s modules described functionality in Fan’s style enhancement embodiment would have been equivalently applicable to the second dialect conversion embodiment as well.” Pet. at 20. There is no explanation as to why a PHOSITA at the time of the invention would have found this to be true.

In Fan’s style separation embodiment, the terminal obtains a first voice and a third voice, extracts a PPG feature from the first voice, extracts a style feature from the third voice using a style separation model, and inputs the style feature and PPG feature into a voice fusion model. Ex1006 at ¶¶[0008]–[0009], [0076]– [0094]. The style separation model extracts timbre and sub-features such as prosody, accent, and speaking speed, and the multihead attention structure outputs a style

embedding vector, for example a 256-dimensional style embedding vector. Ex1006 at ¶¶[0087]–[0095].

In [Fan](#)’s dialect conversion embodiment, the system uses a dialect conversion network. The first voice is a voice of a first dialect, and the second voice is a voice of a second dialect. Ex1006 at ¶¶[0016]–[0019], [0068]–[0071]. [Fan](#) further discloses that the dialect conversion network may include multiple dialect conversion models and may preserve style by inputting a style feature of the first voice along with the feature information of the first voice. Ex1006 at ¶¶[0017]–[0019], [0127], [0162], [0166]–[0167].

Petitioner does not explain why a PHOSITA would take the style separation model and style embedding pathway from [Fan](#)’s style-conversion embodiment, combine it with [Fan](#)’s dialect-conversion pathway, and then graft the result onto Feinauer. Those operations are materially different: one extracts a target style from a third voice; the other converts dialect and may preserve source style. Petitioner proffers no evidence that the style separation model would be able to function in its acoustic extraction role within the dialect conversion embodiment architecture, as would be necessary to meet the limitation of the ’457 Patent. The Petition’s assertion that the two embodiments are “equivalently applicable” is both misleading and inaccurate. Pet. at 20. It would not have been obvious to a PHOSITA at the time of the invention to combine the two embodiments.

Petitioner’s unsupported cross-embodiment reconstruction renders Ground 1 fatally flawed.

6. Petitioner Fails to Show the Claimed Chunking Combination

Claim 1 requires “dividing the speech audio signal into a stream of input chunks,” where each input chunk includes “a pre-defined number of frames of the speech audio signal.” Ex1001 at Cl. 1[a]. Petitioner relies on [Prabhavalkar](#) for this limitation because Feinauer does not disclose the claimed chunking structure. Pet. at 15-18. Petitioner’s theory is that a PHOSITA would have modified Feinauer to use [Prabhavalkar](#)’s chunking through routine reprogramming. *Id.* at 17-18. That assertion is conclusory and **would not result in the claimed invention.**

Feinauer is not a chunk-based ASR system like [Prabhavalkar](#). Rather, it discloses an accent and dialect modification system. Ex1004 at ¶¶[0047]–[0057]. In contrast, [Prabhavalkar](#)’s use of chunks is integral to its ASR architecture, where speech frames are grouped into chunks and processed by a speech-recognition model to predict word elements for transcription. Ex1005, 6:56–7:4, 15:16–22.

As acknowledged by the Petition, “Feinauer provides no further detail regarding how its input unmodulated voice signal is divided into [] ‘portions’ for processing.” Pet. at 15. Insufficient evidence is supplied by the Petition to warrant the conclusion that it would have been obvious to a PHOSITA at the time of the

invention to implement the chunking of **Prabhavalkar** in the system architecture of Feinauer. The Petition neglects to address whether the methods applied in speech recognition would have been compatible with the accent and dialect modification architecture. In addition, **Prabhavalkar** teaches using 25 ms or 30 ms frames along with a particular frame that is “stacked with 3 frames to the left” (Ex1005, 15:16-20), but there is no indication that this grouping of frames and/or window size would have been typical, or even compatible, with accent and dialect modification systems.

Additionally, with respect to the Petition’s suggestion that Feinauer contains a latency problem that **Prabhavalkar**’s chunking would solve, this alleged issue is unfounded. In support of the proposed combination of Feinauer and **Prabhavalkar**, the Petition asserts that a PHOSITA would have been motivated to combine the references to address Feinauer’s alleged latency problem: “A PHOSITA would have been motivated by Prabhavalkar’s express statement that a chunk-based approach may ‘limit the amount of latency incurred,’ increasing speed and performance compared to the unmodified Feinauer method.” Pet. at 17.

However, Feinauer does not describe an unresolved latency “problem”: it describes an intentional latency design. Feinauer expressly states that “the conversation can be broadcast on a lag of which the duration must be long enough for a word to be expressed and analyzed for its accent and dialect strength and then

modified to be expressed by the modification software.” Ex1004 at ¶[0086]. **Prabhavalkar**’s fixed 150ms chunks address an entirely different problem, a design oriented toward speech recognition. Ex1005, Abstract, 20:16-18. **Prabhavalkar**’s goal is to avoid waiting for a complete utterance before running a decoder, (Ex1005, 6:14-17) but Feinauer incorporates no decoder, and as a result, the latency issue addressed by **Prabhavalkar** is not present in Feinauer. Accordingly, given that Feinauer operates with an intentional latency target, there is no latency “problem” in the “unmodified Feinauer method” for **Prabhavalkar**’s chunking to solve.

Furthermore, Feinauer itself acknowledges that its design decreases latency:

One benefit of having the components located on a local machine installed with the software application of the accent and dialect modifier is decreasing latency and saving bandwidth (for example, bandwidth would be required to upload information between a computing device and a cloud server to load components and required models).

Ex1004 at ¶[0081]. Given that the Petition relies on the reduction of latency as the motivation for combining the chunking of **Prabhavalkar** with the system of Feinauer, the Petition’s analysis is fatally flawed. Feinauer already concedes that latency is decreased, and as a result, there is no legitimate reason to combine the references in the proposed manner.

7. Ground 1 Dependent-Claim Challenges Do Not Cure the Independent-Claim Defects

Ground 1 also challenges independent claim 14, along with dependent Claims 2–4, 8, and 15–17. Those dependent-claim offerings do not cure the deficiencies in Petitioner’s showing for the independent claims.

Claims 2–4 and 8 depend from Claim 1. Even if Petitioner could identify isolated teachings for a latency range, a generated speaker embedding, a pretrained embedding, or preceding frames as context, those teachings do not establish the required base pipeline of Claim 1. Petitioner must first show the claimed chunked, context-based extraction of acoustic and linguistic features, speaker-embedding integration, linguistic features with a reduced accent, and output-chunk generation. It has not done so. Similarly, Claims 15–17 depend from Claim 14 and inherit the same deficiencies.

Accordingly, Ground 1 fails for all claims challenged in that ground.

B. Grounds 2–7 Fail on the Merits

1. Grounds 2–7 Fail Because They Rely on Ground 1

Grounds 2–7 fail at least because they depend on Ground 1’s defective base combination. Ground 2 adds Accentron for Claims 9–11; Ground 3 adds Singh for Claims 6 and 19; Ground 4 adds Wu for Claim 7; Ground 5 adds Bongiovi for Claims 12 and 13; Ground 6 adds Osuna for Claim 5; and Ground 7 adds Osuna and Wu for Claim 18. Pet. at v-vii.

Because Petitioner has not shown that Ground 1 satisfies independent claims 1, 14, and 20, Petitioner cannot show that the dependent claims are unpatentable in Grounds 2–7. The additional references are used only for additional dependent limitations. They do not cure Ground 1’s failures regarding chunking, extraction, and speaker-embedding integration. The Board can deny Grounds 2–7 because they depend on Ground 1.

2. Ground 2 Fails Because Accentron Does Not Cure the Missing Claimed Pipeline

Ground 2 challenges Claims 9–11 and adds Accentron for a melspectrogram. The Petition does not provide a technically sufficient explanation of how the integration of Accentron would be implemented in practice, particularly how Accentron’s synthesis module would be adapted to operate within the Feinauer system architecture. **Ground 2 does not offer any fix for the failures of Ground 1 discussed above.**

In any event, Accentron’s seq2seq synthesis architecture operates on neural speech representations and generates melspectrograms (Ex1008 at Abstract), while Feinauer is an accent and dialect modifier using translation models, dictionaries, pairing databases, and context selection. Ex1008 at ¶¶[0047]–[0057], [0061]–[0078]. Petitioner does not explain how Accentron’s synthesis module would receive the claimed inputs from the proposed Ground 1 system, how its

representations would map to Feinauer's architecture, or how the modified system would generate melspectrogram chunks using the claimed contexts.

The Petition asserts that the combination would require only routine reprogramming of the Feinauer-Prabhavalkar-Fan software. Pet. at 48. This is a drastic oversimplification in light of the substantial architectural differences between the systems, as the **claimed invention would not result**. Feinauer's accent modifier 318 is based on symbolic word-level substitution: it identifies discrete words, applies target-accent mappings, and outputs modified audio. Ex1004 at ¶¶[0014], [0036], [0038], Fig. 6. By contrast, Accentron's seq2seq module takes a numerical representation of the source speaker's voice, passes it through neural network layers that reshape its frequency patterns toward the target accent, and outputs a frame-by-frame numerical representation of the target-accented voice (melspectrogram), which a second neural network then converts into an audio waveform. Ex1008, at 1, 3, 12. The numerical representation of the seq2seq module is formed by concatenating bottleneck features (BNFs) with accent-stripped linguistic content, an accent embedding representing the source accent, and a speaker embedding representing the target voice. *Id.* at pp. 2-3, 11-12. This is provided to a decoder that generates the melspectrogram, which is then converted to audio by a vocoder. *Id.* at p. 3. Therefore, these systems operate on

fundamentally different paradigms, such that implementing one within the other would require extensive architectural redesign, not mere routine reprogramming.

In the proposed combination, Accentron's BNFs are substituted for the PPGs of the suggested Feinauer-[Prabhavalkar-Fan](#) model. Pet. at 52. It is critical for Accentron's architecture that the BNFs are accent-independent. Ex1008, at 2, 8. Yet, there is no analysis of whether PPGs, which are reduced in accent but not accent-independent, can function in this role.

Furthermore, the Petition asserts that a PHOSITA would have found it simple to modify Feinauer to produce melspectrograms, as described in Accentron, because Feinauer "already teaches being capable to produce a spectrogram." Pet. at 46. However, Feinauer discloses spectrograms used for analysis on incoming accent (Ex1004 at ¶¶[0064], [0095]), not for use in outputting audio with a converted accent, as disclosed by Accentron. Ex1008 at 8-9.

In accordance, a substitution to include Accentron's neural network seq2seq synthesis module in the Feinauer framework would completely eliminate Feinauer's accent modifier 318 as it originally functioned. This is not a routine reprogramming, and there is no explanation as to how the combination would overcome the above-described constraints. Ground 2 therefore fails.

3. Ground 3 Fails Because Singh's ASR Training Does Not Teach Accent-Reduced Linguistic Feature Extraction

Ground 3 challenges Claims 6 and 19 and adds Singh for a neural network trained using a loss function to reduce accent contributions. Singh is relied on for ASR/domain-adaptation training. Pet. at 57-61. While the Petition suggests implementing Singh's neural network training practices to "increase the accuracy and efficiency of the accent modification" (*id.* at 60), the proposal reflects a flawed cross-domain modification. As with the Feinauer-Prabhavalkar combination, the Petition imports a technique designed exclusively for ASR into Feinauer's accent modification architecture, with the Petition assuming without analysis that the practices disclosed in the ASR context would fit within the accent conversion structure. Pet. at 57-61. The combination **would therefore not result in the claimed invention.** Furthermore, **Ground 3 does not offer any fix for the failures of Ground 1 discussed above.**

In any event, an ASR structure like Singh's lacks the generator components necessary to create audio; it can only identify a category. The Petition provides no sufficient explanation to demonstrate that Singh's training approach could be applied to Feinauer; each system is targeting different outcomes and producing distinct outputs for separate purposes. Ground 3 therefore fails.

4. Ground 6 Fails Because Osuna Does Not Cure the Missing “Linguistic Features With a Reduced Accent” Limitation

Ground 6 challenges Claim 5 and adds Osuna for PPGs with a standardized phonetic dictionary. The Petition does not explain why Feinauer would abandon or modify its existing linguistic representation in favor of the PPGs as taught by Osuna. The Petition asserts that incorporating Osuna would provide the Feinauer-Prabhavalkar-Fan system’s PPGs with “a representation of phonemes”. Pet. at 73. However, Fan already describes PPGs as “equivalent to recognizing the phoneme” and “describ[ing] a probability of each phoneme in a phoneme set corresponding to a voice frame”. Ex1006 at ¶¶[0007]. **Ground 6 does not offer any fix for the failures of Ground 1 discussed above.**

In any event, the Petition’s stated rationale for incorporating Osuna is redundant and **would not result in the claimed invention**. It does not identify any deficiency in Fan’s existing use of PPGs or explain what additional functionality Osuna’s approach would meaningfully contribute beyond what Fan already provides. Nor does it explain why a PHOSITA would replace or modify an already functional phoneme-representing system with Osuna’s alternative framework. Absent such an explanation, the proposed combination relies on hindsight rather than a reasoned basis grounded in the prior art. Ground 6 therefore fails.

5. Ground 5 Fails Because Bongiovi’s Loudness Processing Does Not Cure the Claimed Feature-Extraction and Output-Generation Defects

Ground 5 challenges Claims 12 and 13 and adds Bongiovi for digital signal processing and loudness adjustment. Bongiovi’s digital signal processor (DSP) optimizes audio for consistent loudness and altered dynamic range in playback through output devices in high-noise environments. Ex1010 at ¶¶[0037], [0059]. The Petition argues that “[a] PHOSITA would have appreciated that reducing distortion of Feinauer’s speech audio signal before further processing would have resulted in improved accent modification quality, speed, and computational efficiency by *reducing the amount of unwanted signal being processed.*” Pet. at 67 (italicization added). **Ground 5 does not offer any fix for the failures of Ground 1 discussed above.**

While Bongiovi applies filtering and compression to alter the spectral characteristics of the audio signal (Ex1010 at ¶¶[0038], [0042]-[0043], Fig. 2), it does not remove portions of the speech signal or reduce the amount of data processed; rather, the full signal continues through the processing pipeline with modified amplitude and frequency characteristics. Bongiovi does not eliminate any portion of the signal. The proposed combination **would not result in the claimed invention**, and Ground 5 therefore fails.

6. Ground 4 Fails Because Wu’s VAD Does Not Cure the Missing Acoustic-Feature Extraction Architecture

Ground 4 challenges Claim 7 and adds Wu for the inclusion of a voice activity detector. **Ground 4 does not offer any fix for the failures of Ground 1 discussed above.**

In any event, the Petition neglects to address a critical distinction: in Wu, the VAD is used in the context of automatic gain control (AGC), not accent conversion. Ex1009, Abstract, 3:3-9, Fig. 1. Particularly, Wu employs the VAD to determine whether a given frame contains speech so that an appropriate gain decay rate can be applied. *See id.* 18:1-10. The Petition does not adequately explain why a PHOSITA at the time of the invention would have been motivated to implement Wu’s VAD in the proposed Feinauer system. This combination would not “require[] only reprogramming of *Feinauer*’s software using ordinary skill of a PHOSITA and hardware already present”, as the Petition asserts. Petition, p. 63. The proposed combination **would not result in the claimed invention**, and Ground 4 therefore fails.

7. Ground 7 Fails Because Osuna and Wu Do Not Cure the Defects in Claim 14

Ground 7 challenges Claim 18 and adds Osuna and Wu for PPGs with a standardized phonetic dictionary and the inclusion of a voice activity detector, respectively. As Ground 7 also relies upon the same combinations presented in

Grounds 4 and 6, it fails for the same reasons set forth above with respect to Grounds 4 and 6. Ground 7 therefore fails. Additionally, **Ground 7 does not offer any fix for the failures of Ground 1 discussed above.**

VI. CONCLUSION

Ground 1 is the foundation of the Petition, and Petitioner does not show a supported motivation to combine Feinauer, [Prabhavalkar](#), and [Fan](#) or a reasonable expectation of success. Because Grounds 2–7 depend on Ground 1’s defective base combination, those grounds fail as well. Furthermore, Grounds 2–7 each fail for defective combinations.

For the foregoing reasons, Patent Owner respectfully requests that the Board deny institution of inter partes review of Claims 1–20 of the ’457 Patent.

Respectfully submitted,

AMSTER, ROTHSTEIN & EBENSTEIN LLP
Attorneys for Patent Owner
405 Lexington Avenue, 48th Floor
New York, NY 10174
(212)336-8000

Dated: June 25, 2026
New York, New York

By: /Charles R. Macedo/
Charles R. Macedo (Reg. No. 32,781)

Of counsel:
Gary J. Gershik (Reg. No. 39,992)
Lewis Derenzis III (Reg. No. 84,132)

CERTIFICATE OF WORD COUNT

The undersigned certifies that the foregoing PATENT OWNER'S PRELIMINARY RESPONSE complies with the type-volume limitation in 37 C.F.R. § 42.24(b)(1). According to the word-processing system's word count, the brief contains approximately 10,289 words, excluding the parts exempted by 37 C.F.R. § 42.24(a).

Respectfully submitted,

AMSTER, ROTHSTEIN & EBENSTEIN LLP
Attorneys for Patent Owner
405 Lexington Avenue, 48th Floor
New York, NY 10174
(212)336-8000

Dated: June 25, 2026
New York, New York

By: /Charles R. Macedo/
Charles R. Macedo (Reg. No. 32,781)

CERTIFICATE OF SERVICE

I hereby certify that on June 25, 2026 a true and correct copy of the foregoing PATENT OWNER'S PRELIMINARY RESPONSE was served via email, by consent, to Petitioner's counsel of record at the correspondence addresses of record in this proceeding.

Eric A. Bruesh (Reg. No. 50,394)
Chris R. Schmidt (Reg. No. 63,982)
Nick R. Apel (Reg. No. 84,017)
Adam M. Sandwell (Reg. No. 72,484)
Erise IP, P.A.
7015 College Blvd., Suite 700
Overland Park, Kansas 66211
Telephone: (913) 777-5600
Fax: (913) 777-5601
chris.schmidt@eriseip.com
nick.apel@eriseip.com
adam.sandwell@eriseip.com
PTAB@eriseip.com

Respectfully submitted,

AMSTER, ROTHSTEIN & EBENSTEIN LLP
Attorneys for Patent Owner
405 Lexington Avenue, 48th Floor
New York, NY 10174
(212) 336-8000

Dated: June 25, 2026
New York, New York

By: /Charles R. Macedo/
Charles R. Macedo (Reg. No. 32,781)