

NOTE: This disposition is nonprecedential.

**United States Court of Appeals
for the Federal Circuit**

NUANCE COMMUNICATIONS, INC.,
Appellant

v.

MMODAL LLC,
Appellee

**ANDREW HIRSHFELD, PERFORMING THE
FUNCTIONS AND DUTIES OF THE UNDER
SECRETARY OF COMMERCE FOR
INTELLECTUAL PROPERTY AND DIRECTOR OF
THE UNITED STATES PATENT AND TRADEMARK
OFFICE,**
Intervenor

2020-1549, 2020-1550

Appeals from the United States Patent and Trademark
Office, Patent Trial and Appeal Board in Nos. IPR2018-
01431, IPR2018-01435.

Decided: March 10, 2021

BRIAN E. FERGUSON, Weil, Gotshal & Manges LLP,
Washington, DC, for appellant. Also represented by

STEPHEN BOSCO; ANISH R. DESAI, SUDIP KUMAR KUNDU, New York, NY; DAVID GREENBAUM, Nuance Communication, Inc., Mahwah, NJ.

GABRIEL K. BELL, Latham & Watkins LLP, Washington, DC, for appellee. Also represented by INGE OSMAN, JONATHAN M. STRANG, KEVIN WHEELER; DAVID K. CALLAHAN, Chicago, IL.

MEREDITH HOPE SCHOENFELD, Office of the Solicitor, United States Patent and Trademark Office, Alexandria, VA, for intervenor. Also represented by THOMAS W. KRAUSE, FARHEENA YASMEEN RASHEED, MOLLY R. SILFEN.

Before LOURIE, O'MALLEY, and REYNA, *Circuit Judges*.

LOURIE, *Circuit Judge*.

Nuance Communications, Inc. (“Nuance”) appeals from two final written decisions of the U.S. Patent and Trademark Office Patent Trial and Appeal Board (“the Board”) holding claims 8 and 13 of U.S. Patent 8,117,034 (“the ’034 patent”) and claims 9–11 of U.S. Patent 6,999,933 (“the ’933 patent”) unpatentable as obvious. *See MModal LLC v. Nuance Commc’ns, Inc.*, No. IPR2018-01431 (P.T.A.B. Apr. 3, 2020), J.A. 134–96; *MModal LLC v. Nuance Commc’ns, Inc.*, No. IPR2018-01435 (P.T.A.B. Apr. 21, 2020), J.A. 197–266. For the reasons detailed below, we *affirm*.

BACKGROUND

Nuance owns the ’034 and ’933 patents, which are directed to systems and methods for correcting text generated by automatic speech recognition technology (“ASR”). We begin with a brief background of the technology. ASR converts spoken words into text. J.A. 2702. Specifically, audio files with speech recordings are “distribute[d]” to computers with ASR. ’933 patent col. 1

ll. 23–28.¹ Using ASR, the computers generate a written transcript of the audio file. *Id.* col. 1 ll. 29–39.

The patents describe that ASR can be error-prone, requiring human editors (“transcriptionists”) to make corrections to the converted text. *Id.* col. 1 ll. 4–9; *see also* J.A. 1441 col. 1 ll. 35–52. In order to correct the generated text, transcriptionists typically listen to an audio file of the words while an “audio cursor” follows along in the transcript. ’933 patent col. 1 ll. 40–50. The audio cursor visually indicates the word in the transcript that corresponds to the word that has just been spoken in the audio file. *Id.* This method is referred to as “synchronous playback mode.” *Id.* Although synchronous playback mode made it easier for transcriptionists to review the transcript, it had a specific disadvantage: whenever transcriptionists would spot an error, they would need to stop the playback of the audio, correct the error, and only then resume the audio. *Id.* col. 1 ll. 51–58. The patents explain that the delay could be time consuming. *Id.* col. 2 ll. 7–13.

The patents purport to improve upon the disadvantages of synchronous playback mode. Unlike previous systems, which disclosed only the use of an audio cursor, the patents disclose the use of a synchronous playback mode that includes an audio cursor *and* a text cursor. *Id.* col. 3 ll. 29–52. Consequently, transcriptionists can make a text correction with the text cursor while the audio cursor continues to move through the text in time with the audio. *Id.* col. 6 ll. 35–42. Importantly, transcriptionists need not stop the audio playback when making a text correction, unlike prior systems. *Id.* col. 3 ll. 35–43. The patents further describe that transcriptionists can synchronize the text cursor with the audio cursor or the

¹ Because the specifications of the patent are substantially similar, we cite only the ’933 patent unless otherwise indicated.

audio cursor with the text cursor. *Id.* col. 3 ll. 53–66, col. 8 ll. 1–7. All of the challenged claims recite an audio cursor and a text cursor.

Claim 8 of the '034 patent reads as follows:

8. A method of assisting in correcting text information recognized by a speech recognition device from speech information, the method comprising:

receiving the speech information, the text information recognized from the speech information, and link information that associates portions of the text information with portions of the speech information from which the portions of the text information were recognized by the speech recognition device;

providing an audio cursor for display during acoustic playback of the speech information, the audio cursor highlighting portions of the text information synchronous with the playback of the speech information according to associations provided by the link information such that, when displayed to the user, the audio cursor highlights the portions of the text information as the associated portions of the speech information are being acoustically played back; and

providing a text cursor for display to facilitate editing the text information, the text cursor indicating a position in the text information where at least one edit will be performed upon receiving editing information entered by the user; and

automatically synchronizing the text cursor and the audio cursor, wherein automatically synchronizing the text cursor and the audio cursor comprises automatically positioning the text cursor at a predetermined position relative to a location of

the audio cursor and automatically moving the location of the text cursor synchronous with the movement of the audio cursor during the acoustic playback until an editing operation is performed.

'034 Patent col. 9 l. 43–col. 10 l. 6.

Dependent claim 13 recites:

13. The method of claim 8, wherein automatically synchronizing includes continuously automatically synchronizing the text cursor and the audio cursor when a continuous synchronous playback mode is activated, the method further comprising:

deactivating continuously automatically synchronizing upon receiving at least one first keyboard input from the user, the deactivating including uncoupling the text cursor from the audio cursor; and

activating the continuous synchronous playback mode upon receiving at least one second keyboard input from the user to resume continuously automatically synchronizing the text cursor and the audio cursor.

Id. col. 10 ll. 33–45.

Claim 9 of the '933 patent reads as follows:

9. A correction method (16) for the correction of incorrect words in text information (ETI) recognized by a speech recognition device (1) from speech information (SD), in which the following method steps are executed:

reception of the speech information (SD), the associated recognized text information (ETI) and link information (LI), which marks the part of the speech information (SD) at which the word was

recognized by the speech recognition device (1) for each word of the recognized text information (ETI);

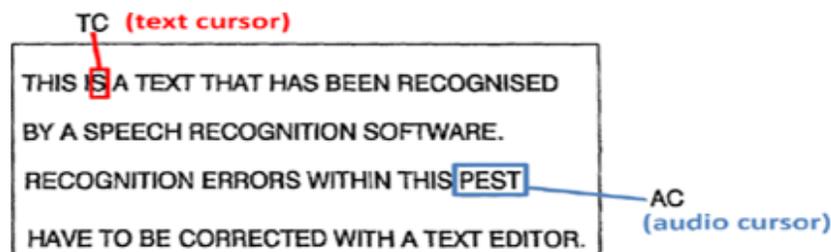
allowing a synchronous playback mode, in which, during the acoustic playback of the speech information (SD) the word of the recognized text information (ETI), which word is marked by the link information (LI) for the speech information (SD) just played back is marked synchronously, while the word just marked features the position of an audio cursor (AC);

editing of the incorrect word with a text cursor (TC) according to editing information (EI) entered by a user, the editing of the incorrect word being possible with the synchronous playback mode activated in the correction device (10).

'933 patent col. 9 l. 44–col. 10 l. 20.

Claim 10, which depends from claim 9, recites that the “text cursor (TC) is synchronized with the audio cursor (AC) or the audio cursor (AC) is synchronized with the text cursor (TC) depending on the editing information entered (EI).” *Id.* col. 10 ll. 21–25. Claim 11, which also depends from claim 9, requires that the “cursors . . . are synchronized by manually actuating at least one key.” *Id.* col. 10 ll. 26–28.

This appeal primarily centers on two elements of the claims: (1) the use of a text cursor to edit incorrect words in a transcript, and (2) the use of an audio cursor to visually indicate the word in the transcript that corresponds to the



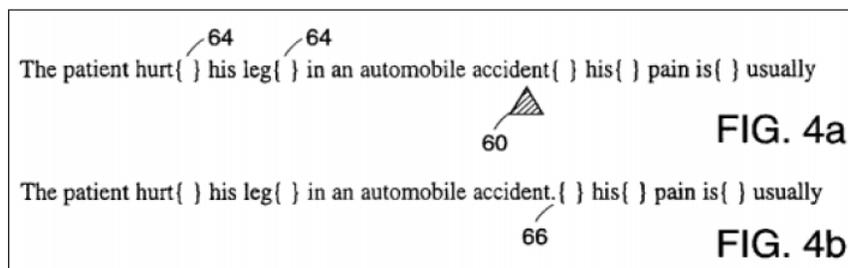
word that has just been spoken in the audio file. See Appellant’s Br. 46. Both cursors are shown below.

’034 Patent, Portion of Fig. 1 (annotated).

MModal LLC (“MModal”) filed petitions for inter partes review of claims 8 and 13 of the ’034 patent and claims 9–11 of the ’933 patent. In both petitions, MModal asserted that the claims would have been obvious over U.S. Patent 6,360,237 (“Schulz”), or Schulz in view of U.S. Patent Publication 2002/0095291 (“Sumner”).

Schulz discloses systems and methods for editing text generated by ASR while the recorded audio is played back. Like the patents, Schulz explains that prior art methods of correcting ASR transcription errors were time consuming because they required the transcriptionists to stop the audio playback before correcting any errors. J.A. 1441 at col. 2 ll. 6–8, 16–24. In order to improve upon this “slow process,” Schulz describes a “playback edit mode” that allows transcriptionists to edit a transcript without stopping the audio recording. J.A. 1441 at col. 2 ll. 23–24; J.A. 1443 at col. 5 l. 54–col. 6 l. 3.

Schulz discusses multiple approaches for defining the location of an edit within the transcript. In one embodiment, one cursor visually indicates *both* the location at which a text edit will occur and the position of the word just spoken in the audio playback. J.A. 1440 at Figs. 4a–4b; J.A. 1444 at col. 7 ll. 29–32; J.A. 1446 at col. 11 ll. 31–36. For example, “[i]n FIG. 4a, [shown below] the cursor is

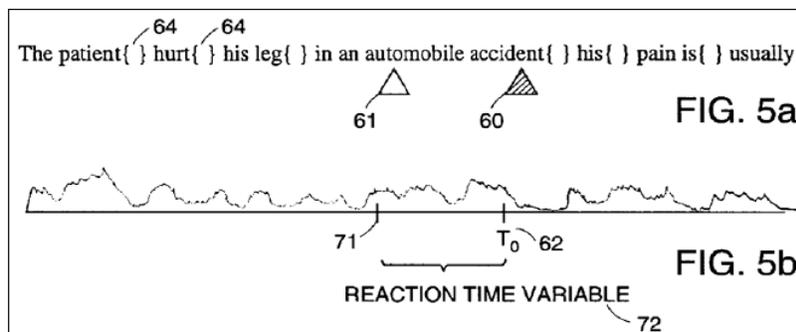


Schulz, Figs. 4a and 4b (Appx1440)

underneath the word ‘accident’ at the same time that this word is being spoken on the audio recording. . . . The period edit function key is then depressed. FIG. 4b shows the insertion of a period 66 immediately after the word ‘accident.’” J.A. 1446 at col. 11 ll. 31–36.

Alternatively, Schulz discloses a second embodiment that includes a “reaction time variable” to improve the editing process. Here, Schulz recognizes that a user may struggle to press the appropriate key quickly enough to trigger an edit “while the desired word is underscored by the cursor.” J.A. 1446 at col. 11 ll. 49–54. The reaction time variable can thus “compensate for the transcriptionist’s reaction time by adjusting the location of an editing function by the reaction time.” *Id.* at col. 11 ll. 63–65. In this embodiment, there are two separate locations: (1) an audio cursor that visually indicates the word in the transcript that corresponds to the word that has just been spoken in the audio file (“cursor 60”), and (2) a text insertion point that is separated from cursor 60 by a period of time and determines the location that text edits will be made (“insertion point 61”). *Id.* at col. 11 l. 49–col. 12 l. 55. *See also* Appellee’s Br. 39. Cursor 60 is denoted by a visual indicator. Important to this appeal, unlike cursor 60, insertion point 61 is not displayed visually (although insertion point 61 is denoted by a triangle in Fig. 5a, the Board found, and the parties do not dispute, that it is technically “not displayed visually”). J.A. 169; J.A. 241.

Figure 5a, shown below, illustrates this embodiment. Schulz explains that “cursor 60 is aligned under the word ‘accident’ as it is being reproduced in audio.” J.A. 1443 at col. 6 ll. 30–35. If a reaction time variable is employed (such as 250 milliseconds), when the user presses the edit function key at time T0, an edit will be performed on the word “automobile” (the word that was marked by the cursor 60 at a time 250 milliseconds before time T0 and is represented by insertion point 61). J.A. 1446 at col. 12 ll. 22–32.



Schulz, Figs. 5a, 5b (Appx1440)

Schulz also discloses a “standard text editor mode,” in which playback of the audio file is stopped and the transcriptionists can use a cursor to edit the text. J.A. 1443 at col. 5 ll. 35–44; J.A. 1447 at col. 13 ll. 21–27.

Sumner discloses two cursors with two different functions for use with ASR: (1) an insertion cursor, to “denote the location where new text will be inserted within a document,” and (2) a correction cursor, which marks the last location where a correction to the text was made. J.A. 1538–40 ¶¶ 8, 14, 24–25.

Before analyzing whether the claims would have been obvious, the Board engaged in claim construction. It determined that “using a *single* visual indication on a display to mark the position of both the audio cursor and the text cursor” falls within the scope of the claims. J.A. 143; J.A. 207 (emphasis added). Subsequently, the Board concluded that the claims are unpatentable as obvious under two separate grounds. Under the first ground, pursuant to its claim construction, the Board determined that the claims would have been obvious over Schulz. Specifically, the Board found that the single displayed cursor in Schulz, as exemplified in Figures 4a–b, satisfies both the “audio cursor” and “text cursor” limitations of the claims. J.A. 165–67; J.A. 228–30. Under the second

ground, the Board concluded that, even if contrary to its construction, the claims require *separate* visual indicators for each cursor, they still would have been obvious over Schulz's reaction time embodiment (Figures 5a–b). Specifically, it determined that it would have been obvious to combine a “visual indicator” at the targeted insertion point 61 (to satisfy the text cursor limitation) with the audio cursor 60, in view of Schulz, or Schulz and Sumner. J.A. 169; J.A. 241.

Nuance appealed to this court. We have jurisdiction pursuant to 28 U.S.C. § 1295(a)(4)(A).

DISCUSSION

We review the Board's legal determinations de novo, *In re Elsner*, 381 F.3d 1125, 1127 (Fed. Cir. 2004) (citing *In re Kollar*, 286 F.3d 1326, 1329 (Fed. Cir. 2002), and its fact findings for substantial evidence, *In re Gartside*, 203 F.3d 1305, 1316 (Fed. Cir. 2000). A finding is supported by substantial evidence if a reasonable mind might accept the evidence as adequate to support the finding. *Consol. Edison Co. v. N.L.R.B.*, 305 U.S. 197, 229 (1938).

Obviousness is a question of law, supported by underlying fact questions. *In re Baxter Int'l, Inc.*, 678 F.3d 1357, 1361 (Fed. Cir. 2012). In evaluating obviousness, we consider the scope and content of the prior art, differences between the prior art and the claims at issue, the level of ordinary skill in the pertinent art, and any relevant secondary considerations. *Graham v. John Deere Co.*, 383 U.S. 1, 17–18 (1966).

Nuance asserts that the Board erred in concluding that the prior art renders obvious the use of a text cursor in combination with an audio cursor, as required by the claims. We first address Nuance's arguments regarding independent claims 8 and 9 and then address Nuance's arguments regarding dependent claims 10 and 13.

I

We turn first to Nuance’s argument that the Board erred in concluding that claims 8 and 9 are unpatentable as obvious. Nuance contends that the Board’s second obviousness ground, which was based on Schulz’s reaction time embodiment, was erroneous. Specifically, Nuance asserts that the Board erred in finding that it would have been obvious for a person of skill to combine a (1) “visual indicator” at insertion point 61 with (2) the audio cursor 60. According to Nuance, the Board provided “no . . . reasons” as to why a person of skill in the art would be motivated to modify Schulz in this manner. Appellant’s Br. 53. MModal responds that the Board’s analysis was supported by substantial evidence. According to MModal, a person of ordinary skill would have been motivated to add a visual indicator at insertion point 61 in order to view the location where text edits would occur.

We agree with MModal that the Board’s determination was supported by substantial evidence. First, the Board found that “the use of, and the benefits of displaying, each type of cursor—an audio cursor and a text cursor—were well known in the art.” J.A. 169; *see also* J.A. 241. As the Board observed, the patent specifications themselves disclose that it was well known in the art to use an audio cursor to follow the words being played back and to use a text cursor to make corrections to the text. *See* J.A. 169 (citing ’034 patent col. 1 ll. 28–56); *see also* J.A. 241. Given these benefits, the Board reasonably found that a person of ordinary skill would have been motivated to add a visual indicator at insertion point 61 for use with Schulz’s audio cursor 60. The Board elaborated that doing so would allow a person to simultaneously (1) confirm the “precise position” where edits would occur with a text cursor at insertion point 61, and (2) observe the text being spoken in the audio playback with the audio cursor 60. J.A. 172–174; J.A. 245–247. Indeed, the Board pointed out that *not* displaying a visual indicator at insertion point 61 could

create confusion as to where the corrections would be made. J.A. 172 (citing expert testimony); J.A. 245. The Board's determination was further supported by Sumner, which discloses the advantages of displaying cursors relating to two different, but relevant functions at the same time. *Id.* Moreover, implementing such a modification would have taken only routine skill. For example, the Board credited MModal's argument, based on expert testimony, that "it was well-known for text editors, such as Microsoft Word, to visually display a text cursor . . ." J.A. 169; J.A. 241–42.

Nuance makes several arguments as to why the Board's determination should be reversed, all unconvincing. First, Nuance argues that Sumner cannot support the Board's second obviousness determination because its two cursors do not function as an audio cursor and a text cursor. Appellant's Br. 53.

Nuance's argument misses the mark. The Board acknowledged that, strictly speaking, Sumner's cursors do not correspond to the audio cursor and text cursor of the claims. *See* J.A. 173–75; J.A. 246–48. However, the Board did not rely on Sumner for that purpose. Rather, the Board relied on Sumner's disclosure that it would be beneficial to simultaneously display two cursors with different functions. *Id.* The Board further found that a person of ordinary skill would have been motivated to implement that teaching in order to modify Schulz. *Id.*

Second, Nuance argues that the Board improperly "requir[ed]" it to prove that the specifications describe the inventive aspects of the claims, namely, the simultaneous display of two separate cursors. Appellant's Br. 51–52. According to Nuance, the specifications "did not need to emphasize the visual nature of the cursors because" a person of ordinary skill in the art "reviewing the specifications would have understood . . . the advantages that the two cursors would provide." *Id.* at 52.

Consequently, Nuance contends it was improper for the Board to require such a showing in order to conclude that the claims are nonobvious. *Id.*

We disagree with Nuance's interpretation of the Board's analysis. Contrary to Nuance's assertion, at no point did the Board indicate that its obviousness determination hinged on the specifications' disclosure that the claimed subject matter is inventive. Rather, the Board simply examined the specifications and found support for its obviousness determination. It was not improper for the Board to review the specifications when analyzing whether the claims would have been obvious. Moreover, the Board's analysis was further supported by a variety of other evidence, including prior art and expert testimony. *See, e.g.,* J.A. 169; J.A. 241–42.

Third, Nuance asserts that insertion point 61 cannot be a “different cursor from cursor 60” as it is “simply the position of cursor 60 at a different, earlier point in time.” Appellant's Br. 49. We disagree. Although insertion point 61 denotes the location of cursor 60 at an earlier point in time, it still performs a separate and distinct function, namely, denoting where edits will occur. *See* J.A. 169. Indeed, for that very reason, in *Schulz*, it is denoted by a different number (61) from cursor 60.²

² Nuance also contends that the Board should have construed cursor to mean a “moveable indicator on a display screen.” Appellant's Br. 35, 52–53. According to Nuance, under its construction, insertion point 61 cannot be a cursor because it doesn't visibly indicate the text location. However, that amounts to an argument that because insertion point 61 is not visible, it would not have been obvious to make it visible. As discussed, the Board already found that it would have been obvious to make insertion point 61 visible.

In sum, the Board reasonably found that a person of ordinary skill “would have wanted, and known how, to see where . . . text edits would occur by providing a visual indicator” to show the position of insertion point 61. J.A. 173; J.A. 245. Because the Board’s second obviousness ground was supported by substantial evidence, we need not reach Nuance’s arguments regarding the Board’s first obviousness ground (Schulz Figures 4a–b), which did not include a reaction time.

II

We turn next to Nuance’s arguments that the Board erred in concluding that dependent claims 10 and 13 are unpatentable as obvious. First, Nuance asserts that the Board did not explicitly determine that claims 10 and 13 would have been obvious under its second obviousness ground (based on Schulz’s reaction time embodiment). Rather, Nuance contends that the Board applied its second obviousness ground only to independent claims 8 and 9. Second, Nuance argues that claims 10 and 13 are not unpatentable as obvious, even under the Board’s second obviousness ground. We address each argument in turn.

A

As an initial matter, the Board indicated that it analyzed the dependent claims under the first *and* second obviousness grounds. For example, when addressing Nuance’s arguments regarding claim 10, the Board clarified that it was incorporating its entire analysis for claim 9, including its second obviousness ground. *See* J.A. 257 (noting with respect to claim 10 that “[f]or reasons we discussed for claim 9 . . . we are persuaded [MModal] has shown Schulz, alone or in combination with Sumner, teaches two distinct cursors”). Similarly, the Board stated that its analysis of claim 13 incorporates its analysis of claim 8, at least in part. *See* J.A. 185 (reiterating that “all of” claim 8’s limitations would have been obvious “[f]or the reasons explained above in Section III.B.3.a” and then

explaining why “the *additional* limitations recited in claim 13” would *also* have been obvious in view of Schulz) (emphasis added).

To the extent that the Board did not explicitly state whether it was analyzing the claims under the second obviousness ground, any such ambiguity was harmless. First, the logic of the Board’s determination would have been substantially the same under either obviousness ground, as will be further explained below. Second, the dependent claims only add minor limitations to the independent claims. For example, independent claim 9 recites use of a text cursor and an audio cursor. *See* ’933 patent col. 9 l. 44–col. 10 l. 20. Claim 10, which depends from claim 9, recites, in relevant part, that the “text cursor (TC) is synchronized with the audio cursor (AC) or the audio cursor (AC) is synchronized with the text cursor (TC) depending on the editing information entered (EI).” *Id.* col. 10 ll. 21–25. Indeed, before the Board, Nuance did not present arguments for claim 10 beyond those already raised for claim 9. *See* J.A. 257. Additionally, claim 10’s synchronization limitation is similar to independent claim 8’s limitation that requires “automatically synchronizing the text cursor and the audio cursor,” which the Board explicitly found obvious under its second ground. *Compare* ’933 patent col. 10 ll. 21–25, *with* ’034 patent col. 9 ll. 66–67; J.A. 173–77. Similarly, claim 13 recites, in relevant part, “uncoupling the text cursor from the audio cursor.” ’034 patent col. 10 ll. 40–41.

B

We now turn to Nuance’s argument that claims 10 and 13 are not unpatentable as obvious, even under the Board’s second obviousness ground. With respect to claim 10, Nuance contends that the determination whether to synchronize the text cursor to the audio cursor, or vice-versa, depends on the editing information entered by the user. Nuance asserts, however, that in Schulz, “the

positions of cursor 60 and location 61 (the edit insertion point) are always determined solely by the location of the audio cursor 60 because the location 61 is merely the position of audio cursor 60 at an earlier point in time.” Appellant’s Br. 56–57. We are unpersuaded by Nuance’s argument. As the Board determined, for claim 10, Schulz’s insertion point 61 (the text cursor) realigns (i.e., synchronizes) with the audio cursor after the user *presses an edit function key*. See J.A. 256–57. Moreover, this process remains the same regardless whether there is a reaction time variable (as in the second obviousness ground) or there is no reaction time variable (as in the first obviousness ground).

With respect to claim 13, Nuance asserts that the claim language requires decoupling the text and audio cursors, whereas the edit insertion point 61 in Schulz is “always tied to the cursor 60.” Appellant’s Br. 55. We disagree. Here, the Board found that Schulz’s insertion point 61 is coupled to the audio cursor 60 *only* in synchronous playback mode. J.A. 187–88. The Board then reasonably determined, with support from the specification, that when playback mode is stopped and changed to text editor mode, the audio and text cursors can be uncoupled such that the text cursor can be used as a normal text editor. See, e.g., J.A. 188 (quoting J.A. 1447 at col. 13 ll. 22–27). Moreover, this process remains the same regardless whether there is a reaction time variable (as in the Board’s second obviousness ground) or there is no reaction time variable (as in the Board’s first obviousness ground).³

Nuance also asserts that this court’s decision in *Arthrex v. Smith & Nephew, Inc.*, 941 F.3d 1320 (Fed. Cir.

³ To the extent that Nuance’s arguments here are duplicative of its arguments regarding the independent claims, our analysis regarding the independent claims applies here too.

NUANCE COMMUNICATIONS, INC. v. MMODAL LLC

17

2019), which issued before the final written decisions at issue here, did not cure the Appointments clause defect. However, we have reiterated that final written decisions issued after *Arthrex* were decided by constitutionally appointed Administrative Patent Judges. *See Caterpillar Paving Prods. Inc. v. Wirtgen Am., Inc.*, 957 F.3d 1342, 1343 (Fed. Cir. 2020); *see also Infineum USA L.P. v. Chevron Oronite Co. LLC*, No. 2020-1333, 2021 WL 210722, at *8 (Fed. Cir. Jan. 21, 2021).

CONCLUSION

We have considered Nuance's remaining arguments and find them unpersuasive. For the foregoing reasons, the decisions of the Board are *affirmed*.

AFFIRMED