

Internal distribution code:

- (A) [-] Publication in OJ
- (B) [-] To Chairmen and Members
- (C) [-] To Chairmen
- (D) [X] No distribution

**Datasheet for the decision
of 8 August 2022**

Case Number: T 1206/19 - 3.5.03

Application Number: 14193441.4

Publication Number: 2876903

IPC: H04R25/00

Language of the proceedings: EN

Title of invention:

Spatial filter bank for hearing system

Patent Proprietor:

Oticon A/S

Opponent:

Sivantos Pte. Ltd.

Headword:

Spatial filterbank/OTICON

Relevant legal provisions:

EPC Art. 56

Keyword:

Inventive step - main and auxiliary requests 1 to 3 (no) -
auxiliary request 4 (yes)

Unadmitting admitted evidence - prior-art document D17 (no):
"unadmitting" would prevent a judicial review of the merits of
the first-instance decision - T 960/15 not followed

Decisions cited:

G 0007/93, T 0039/93, T 0960/15



Beschwerdekammern
Boards of Appeal
Chambres de recours

Boards of Appeal of the
European Patent Office
Richard-Reitzner-Allee 8
85540 Haar
GERMANY
Tel. +49 (0)89 2399-0
Fax +49 (0)89 2399-4465

Case Number: T 1206/19 - 3.5.03

D E C I S I O N
of Technical Board of Appeal 3.5.03
of 8 August 2022

Appellant I: Oticon A/S
(Patent Proprietor) Kongebakken 9
2765 Smørum (DK)

Representative: Cohausz & Florack
Patent- & Rechtsanwälte
Partnerschaftsgesellschaft mbB
Bleichstraße 14
40211 Düsseldorf (DE)

Appellant II: Sivantos Pte. Ltd.
(Opponent) 18 Tai Seng Street
No. 08-08
18 Tai Seng
Singapore 539775 (SG)

Representative: FDST Patentanwälte
Nordostpark 16
90411 Nürnberg (DE)

Decision under appeal: **Interlocutory decision of the Opposition
Division of the European Patent Office posted on
18 February 2019 concerning maintenance of the
European Patent No. 2876903 in amended form.**

Composition of the Board:

Chair K. Bengi-Akyürek
Members: K. Schenkel
C. Heath

Summary of Facts and Submissions

- I. The appeals by the patent proprietor and the opponent lie from the interlocutory decision of the opposition division maintaining the present European patent in amended form on the basis of a "third auxiliary request" filed during the oral proceedings before the opposition division on 14 December 2018. The claim request as maintained is identical to auxiliary request 1 underlying these appeal proceedings.
- II. In this decision, reference is made to the following prior-art documents:
- D1:** US 2012/0020505 A1;
D8: EP 2 640 094 A1;
D17: US 6 349 278 B1.
- III. Oral proceedings were held before the board on 8 August 2022 by videoconference.

The parties' final requests were as follows:

- The appellant/opponent requests that the decision under appeal be set aside and that the patent be revoked.
- The appellant/patentee requests that the decision under appeal be set aside and that the patent be maintained as granted (**main request**), in the auxiliary, that the appellant/opponent's appeal be dismissed (**auxiliary request 1**), or further in the auxiliary, that the decision under appeal be set aside and that the patent be maintained on the

basis of one of **auxiliary requests 2 to 10** filed with letter dated 14 November 2019.

At the end of the oral proceedings, the board's decision was announced.

IV. Claim 1 of the **main request** reads as follows (board's labelling):

"A hearing system (10) configured to be worn by a user (62), which comprises, an environment sound input unit (12, 14), an output transducer (18), and electric circuitry (16), wherein the environment sound input unit (12, 14) is configured to receive sound (20) from the environment of the environment sound input unit (12, 14) and to generate sound signals (22, 24) representing the environment, wherein the output transducer (18) is configured to stimulate hearing of a user (62), wherein the electric circuitry (16) comprises a spatial filterbank (34), and wherein the spatial filterbank (34) is configured to use the sound signals (22, 24) to generate spatial sound signals (56) dividing a total space (60) of the environment sound (20) in a plurality of subspaces (58), defining a configuration of subspaces, and wherein a spatial sound signal (56) represents sound (20) coming from a subspace (58),

CHARACTERIZED IN THAT

the electric circuitry (16) comprises a voice activity detection unit (38) configured to determine whether a voice signal is present in a respective spatial sound signal (56) and configured to run voice activity detection in parallel in the different subspaces in a continuous mode, where the voice activity detection unit (38) is configured to estimate a probability for

the voice signal to be present in the spatial sound signal."

- V. Claim 1 of **auxiliary requests 1 and 2** differs from claim 1 of the main request in that the following feature has been added after "CHARACTERIZED IN THAT":

"the electric circuitry (16) comprises a user control interface (50) configured to allow a user (62) to adjust the configuration of subspaces (58), and".

- VI. Claim 1 of **auxiliary requests 3 and 4** differs from claim 1 of auxiliary request 1 in that the added feature now reads as follows (board's highlighting of amended text):

"the electric circuitry (16) comprises a user control interface (50) configured to allow a user (62) to adjust the configuration of subspaces (58) for selecting to listen to the output of a single spatial sound signal (56), and".

- VII. Claim 17 of **auxiliary request 3** reads as follows (board's labelling):

"A method for processing sound signals (22, 24) representing sound (20) of an environment by means of an electric circuitry (16), the method comprising the steps:

- receiving sound signals (22, 24) representing sound (20) of an environment,
- using the sound signals (22, 24) to generate spatial sound signals (56), wherein each spatial sound signal (56) represents sound (20) coming from

a subspace (58) of a total space (60) of the environment sound (20), wherein the electric circuitry (16) comprises a user control interface (50) configured to allow a user (62) to adjust the configuration of subspaces (58) for selecting to listen to the output of a single spatial sound signal (56), and,

CHARACTERIZED IN THAT

it comprises the steps of:

- detecting whether a voice signal is present in the selected single spatial sound signal (56) by running voice activity detection in parallel in the different subspaces in a continuous mode, where the voice activity detection unit (38) is configured to estimate a probability for the voice signal to be present in the spatial sound signal,
- selecting the single spatial sound signal (56) with a voice signal above a predetermined signal-to-noise ratio threshold,
- generating an output sound signal (28) from the selected spatial sound signal (56)."

VIII. Claim 17 of **auxiliary request 4** reads as follows (board's highlighting of amendments vis-à-vis claim 17 of auxiliary request 3):

"A method for processing sound signals (22, 24) representing sound (20) of an environment by means of an electric circuitry (16), wherein the electric circuitry (16) comprises a user control interface (50) configured to allow a user (62) to adjust the configuration of subspaces (58) for selecting to listen

to the output of a single spatial sound signal (56),
the method comprising the steps:

- receiving sound signals (22, 24) representing sound (20) of an environment,
- using the sound signals (22, 24) to generate spatial sound signals (56), wherein each spatial sound signal (56) represents sound (20) coming from a subspace (58) of a total space (60) of the environment sound (20),

~~wherein the electric circuitry (16) comprises a user control interface (50) configured to allow a user (62) to adjust the configuration of subspaces (58) for selecting to listen to the output of a single spatial sound signal (56), and,~~

CHARACTERIZED IN THAT

it comprises the steps of:

- allowing a user (62) to adjust the configuration of subspaces (58) for selecting to listen to the output of a single spatial sound signal (56),
- detecting whether a voice signal is present in the selected single spatial sound signal (56) by running voice activity detection in parallel in the different subspaces in a continuous mode, where the voice activity detection unit (38) is configured to estimate a probability for the voice signal to be present in the spatial sound signal,
- selecting the single spatial sound signal (56) with a voice signal above a predetermined signal-to-noise ratio threshold,
- generating an output sound signal (28) from the selected spatial sound signal (56)."

Reasons for the Decision

1. Technical background of the patent

The opposed patent relates to a hearing system comprising a sound input unit producing sound signals, circuitry to process the sound signals and an output transducer to output the processed sound signals. The circuitry includes a "spatial filterbank" which generates spatial sound signals which separate the environment sound into subspaces (i.e. different directions). The circuitry further includes a "voice activity detection unit" configured to determine whether a voice signal is present in the spatial sound signals of the different subspaces, wherein the voice activity detection unit is further configured to estimate a probability of a presence of a voice signal in the spatial sound signal.

2. *Main request - inventive step (Article 56 EPC)*

2.1 Claim 1 as granted (**main request**) comprises the following limiting features:

- (a) A hearing system configured to be worn by a user, which comprises an environment sound input unit, an output transducer and electric circuitry,
- (b) wherein the environment sound input unit is configured to receive sound from the environment of the environment sound input unit and to generate sound signals representing the environment,
- (c) wherein the output transducer is configured to stimulate hearing of a user,
- (d) wherein the electric circuitry comprises a spatial filterbank which is configured to use the sound signals to generate spatial sound signals dividing

- a total space of the environment sound in a plurality of subspaces, defining a configuration of subspaces, and wherein a spatial sound signal represents sound coming from a subspace,
- (e) wherein the electric circuitry comprises a voice activity detection unit configured to determine whether a voice signal is present in a respective spatial sound signal and configured to run voice activity detection in parallel in the different subspaces in a continuous mode,
 - (f) where the voice activity detection unit is configured to estimate a probability for the voice signal to be present in the spatial sound signal.

Feature (a) defines the hearing system as "configured to be worn by a user" without providing further details as to how exactly this is achieved. Since to "wear" in a broader sense also can mean simply to "carry", the board interprets this feature accordingly in that breadth.

- 2.2 Document **D1** is taken as closest prior art and it also refers to a method and an apparatus for the detection of speech or voice in a mixed sound signal comprising a plurality of excitations (abstract).

As to **features (a) to (c)**, document D1 discloses a hearing system ("hearing aid") with an environment sound input unit ("multiple microphones"), an electric circuitry and an output transducer ("earphones"; see paragraphs [0073] and [0074]). In one embodiment of D1, the hearing aid body containing microphones is placed on a table during its use, whereas the earphones are worn by the user (paragraph [0077], Fig. 6). In a further embodiment, a microphone is worn by the user (paragraph [0191]). The microphones capture

environmental sound signals and convert it into sound signals (paragraph [0073] and FIG. 5). The earphones emit a sound signal and are configured accordingly.

As to **feature (d)**, the hearing aid comprises a spatial filterbank ("excitation separation section") which separates the sound signal by direction, i.e. into subspaces, or, in other words, generates k spatial sound signals (paragraph [0080]).

As to **feature (e)**, the circuitry of the hearing aid includes a voice activity detection unit ("speech detection section") which is configured to determine whether there is a voice signal present in each of the spatial sound signals (see paragraphs [0082] and [0083]: "Next, the processing in step S130 is performed on sound signals S_k ..."; "speech detection section 140 performs speech/non-speech detection on sound signal S_k "). If speech is detected, the corresponding time section is defined as an "utterance section" and a "degree of establishment of a conversation" between sound signals from different directions is determined by analysing the time of overlap of utterance and silence intervals of two different subspaces (paragraphs [0085], [0087] and [0101]). The purpose is to detect whether there is an ongoing conversation between speech sources in two different subspaces. Furthermore, the continuous monitoring for conversations between subspaces inevitably implies a voice activity detection in parallel for all subspaces in a continuous mode. More specifically, for a given time interval, the ratio of the duration of the utterance sections in a first subspace and the duration of utterance sections of a first and a second subspace is calculated (paragraph [0151]). The board also notes

that whether or not a section is defined as "utterance section" is binary in D1.

2.3 Hence, the hearing system of claim 1 differs from the hearing system of D1 in that the "voice activity detection unit" is configured to estimate a probability for the voice signal to be present in the spatial sound signal (i.e. **feature (f)** of claim 1).

2.4 It is conspicuous to the board that the appealed decision is entirely silent as to the formulation of a proper objective technical problem in the framework of the problem-solution approach. The board finds that the technical effect of the distinguishing feature is that the result of the voice detection can reflect uncertain situations in which a binary speech/non-speech detection could produce significant errors. The resulting objective technical problem underlying the invention can thus be seen in "improving the accuracy of the speech detection performed in the system of D1". The problem framed by the proprietor, i.e. "improving spatial processing" in general is considered to be too broad since there are typically countless options for improving spatial processing in a system such as that of D1.

2.5 Prior-art document **D17** has been filed by the opponent with its letter dated 12 October 2018 in response to the patent proprietor's reply to the notice of opposition and an annex to the summons to oral proceedings issued on 17 May 2018. The proprietor argued that D17 should not be admitted into the proceedings, since D17 provided similar information as document **D15** filed by the proprietor and that D17 was not *prima facie* relevant as it was not directed to a

hearing aid.

The board holds that there were indeed good reasons to file document D17 after the opposition division did not share the opponent's line of arguments. As to the similarity to D15, it is noted that the title of D15 refers to a posteriori speech presence probability estimation whereas D17 has no such limitation. The board also agrees with the findings in point 2.3 of the appealed decision and concludes that the opposition division had exercised its discretion to admit D17 into the proceedings without having committed any procedural violation, let alone a substantial one. Thus, the board cannot see any reason to "unadmit" a document which had already been admitted by the first-instance department without the latter having committed any substantial procedural violation in that regard (which would normally lead to a direct remittal of the case). At any rate, as held in **T 39/93** (cf. Reasons 3.1.1, second paragraph), no valid judgment on the *merits* of the first-instance's decision could be made if evidence that was admitted by the first-instance department would simply be "unadmitted" by a board (contrary to the conclusions drawn e.g. in **T 960/15**, Reasons 3, applying the test proposed in an *obiter dictum* of **G 7/93**, Reasons 2.6, relating to the exercise of discretion under Rule 86(3) EPC 1973 (Rule 137(3) EPC).

- 2.6 When trying to solve the above objective problem, the skilled person would have considered document **D17** which refers to a method for providing a speech-probability estimate indicating the probability that a signal includes a speech signal (cf. abstract). The board agrees with the opponent that, when starting out from D1 and faced with the objective problem, the skilled person in the field of hearing aids would have modified

the "voice activity detection unit" according to the teaching of D17 such that it estimates a probability for the voice signal to be present in each spatial sound signal and would have arrived at a system with all the features of claim 1 without exercising inventive skills.

2.7 The proprietor argued that D1 in paragraphs [0083] to [0085], [0101], [0105] and [0110] clearly disclosed that *binary* decisions were made for the speech detection and, once speech was detected, the corresponding time interval was defined as utterance section. It was then looked at the utterance overlap to determine which is the most likely conversation partner on the basis of the overlap of utterances. D1 used "hard decisions" to determine whether or not there were an utterance and to determine the probability for a conversation between two spatial sound signals or subspaces. Furthermore, D1 could not use a speech probability for determining the conversation probability since it required hard decisions and it would further require a complete reconfiguration of the system of D1. Hence, the skilled person could not use the teaching of D17 since it was not combinable with D1. Conversely, the patent determined the probability of speech being present (see paragraph [0049]). The associated objective problem to be solved was therefore "to improve spatial processing".

2.8 The board does not agree with the proprietor's arguments. Document D1 calculates durations, i.e. time lengths, in which the utterance time intervals overlaps or coincides for detecting conversations (see e.g. paragraphs [0087] to [0097]). In order to take into account the speech probability, each time unit of the utterance time intervals could be weighted with the

speech probability for that time unit. The result would be the same as that obtained before a duration reflecting the probability for a conversation.

As to document D17, the board acknowledges that, in its background section, it refers to mobile communication systems and does not mention hearing aids (column 1, lines 17 and 18). However, D17 does not limit the field of application in the claims or the description. Thus, the disclosed method is applicable to every system which processes sound signals including speech signals. Further, the skilled person in the field of hearing systems would in general consult documents relating to speech processing such as D17. The teaching of document D1 thus can indeed be combined with D17 along with minor adaptations. The method of D1 does not necessarily require "hard decisions" as to speech or non-speech, but can also detect conversations based on "soft speech/non-speech decisions", thereby even improving the accuracy. The board also notes in that regard that the probability calculated in D1 refers to the presence of a conversation between subspaces and is independent of a further probability calculated for detecting speech.

2.9 Consequently, Article 100(a) in conjunction with Article 56 EPC prejudices the maintenance of the patent as granted.

3. *Auxiliary requests 1 and 2 - inventive step (Article 56 EPC)*

3.1 Claim 1 of **auxiliary requests 1 and 2** add to the hearing system that

(g) a user control interface is configured to allow a user to adjust the configuration of subspaces.

3.2 No further details as to the exact parameters of the configuration to be adjusted are given in feature (g). So, as put forward by the opponent, even switching between a *static* and an *adaptive* operation mode (see e.g. paragraph [0060] of the patent) would fall within the broad terms of "configuration of subspaces". Hence, feature (g) solely provides the user with the possibility of adjusting such parameters. As a result, the board accepts that the additional (partial) objective problem associated with feature (g) resides in "allowing a user to influence the spatial processing of sounds in an acoustically challenging environment". This problem is however unrelated to the (partial) objective problem underlying feature (f) of claim 1 (see point 2.4 above).

3.3 Document D1 already mentions some generally adjustable parameters related to the subspaces, for example the number of the subspaces ("directions") or how many of them are considered for detecting conversations (see paragraphs [0080] and [0108]). Moreover, to envisage a user control for providing the possibility of adjusting parameters to the user's need were well-known to the skilled person at the patent's filing date. In addition, technical obstacles to applying a user control in the field of hearing systems are not present, as can be seen e.g. in document **D8** relating to remote-controllable hearing systems (paragraph [0002], last sentence).

3.4 The proprietor submitted that the control interface enabled the user to adjust the system and the spatial filterbank to its needs and to improve subjectively the

spatial processing (referring to paragraph [0021] of the patent). The objective technical problem thus was again to *improve spatial processing*. But document D1 did not include any hint that the number of the subspaces is variable or to modify the "configuration of the subspaces". Rather, each hearing aid had a fixed set.

The board holds however that the effects of the adjustment of the "configuration of subspaces" are not limited to *improving the spatial processing*. Further, the parameters related to the configuration of subspaces mentioned in D1, i.e. the number of subspaces and how many of them are considered for the conversation detection, are by their nature freely adjustable and have an effect that is directly noticeable by the user. No specific hint is thus required for a skilled person to provide a user with a possibility of adjusting those parameters in order to make the system more adaptive.

3.5 The board therefore concludes that providing a user merely with the possibility of controlling adjustable parameters in order to make the system of D1 more adaptive cannot contribute to an inventive step.

3.6 Consequently, auxiliary requests 1 and 2 are not allowable under Article 56 EPC.

4. *Auxiliary request 3 - inventive step (Article 56 EPC)*

4.1 Claim 1 of **auxiliary request 3** includes the further limitation that

(h) the user control interface is configured to allow the user to select to listen to the output of a single spatial sound signal.

4.2 Added feature (h) allows the user to steer the attention into a specific direction (i.e. subspace). As indicated in paragraph [0061] of the opposed patent, feature (h) enables the user to listen to another subspace than e.g. a *frontal* subspace, i.e. to speedily switch to sound coming from *non-frontal* directions like in a car-cabin scenario. The objective technical problem underlying feature (h) can thus be seen in "making the hearing system better adaptable for the user to the specific acoustic environment".

4.3 Document D1 does not provide any hint towards selecting a single spatial signal to listen to. To the contrary, D1 teaches to monitor several spatial signals to detect a conversation. D1 does also not teach to steer the sensitivity of the hearing system under the user's control, let alone by providing the user with the possibility of selecting a single spatial signal. No further counter-arguments in that regard were advanced by the opponent at the oral proceedings before the board.

4.4 Hence, the skilled person, starting out from D1, faced with the above objective problem and considering D17 would not have arrived at a hearing system with all the features of claim 1 without exercising inventive skills (Article 56 EPC).

4.5 As to independent **claim 17** of auxiliary request 3, it includes the following limiting features:

- A) A method for processing sound signals representing sound of an environment by means of an electric circuitry, the method comprising the steps:
- B) receiving sound signals representing sound of an environment,
- C) using the sound signals to generate spatial sound signals, wherein each spatial sound signal represents sound coming from a subspace of a total space of the environment sound,
- D) wherein the electric circuitry comprises a user control interface configured to allow a user to adjust the configuration of subspaces for selecting to listen to the output of a single spatial sound signal,
- E) detecting whether a voice signal is present in the selected spatial sound signal by running voice activity detection in parallel in the different subspaces in a continuous mode, where the voice activity detection unit is configured to estimate a probability for the voice signal to be present in the spatial sound signal,
- F) selecting the single spatial sound signals with a voice signal above a predetermined signal-to-noise ratio threshold,
- G) generating an output sound signal from the selected spatial sound signals.

4.6 Feature (D) mirroring feature (h) of claim 1 only specifies the "electric circuitry" for carrying out the claimed method but does not limit the actual method steps themselves. In other words, this feature would, at best, merely contribute to solving the problem of "how to ensure that known or obvious method steps (B), (C) and (E) to (G) are executed by an apparatus with a user interface according to feature (h) or (D). Feature (D) of claim 17 alone can however not

contribute to an inventive step. In addition, features (E), (F) and (G) refer to a "selected" signal without specifying *how* the selected signal is actually chosen. Hence, the selected signal is just the signal which is processed or simply a signal. The reference to a selected signal does therefore not really limit the method.

- 4.7 In sum, the method of claim 17 differs from that of D1 only in that spatial sound signals including a voice signal *above* a predetermined signal-to-noise ratio threshold are selected, i.e. processed (i.e. **feature (F)** of claim 17).
- 4.8 This distinguishing feature ensures a minimum acceptable quality of the sound signals used for the generation of the output signal and thereby improves the intelligibility of the output signal. However, feature (F) together with its effect is not related to (i.e. independent from) the features in claim 1 of the higher-ranking claim requests, and its contribution to an inventive step can therefore be treated separately (juxtaposition).
- 4.9 It is apparent to the board that document D1 already teaches that a high signal-to-noise-ratio may be detrimental (see paragraph [0114]). Excluding the worst signals, i.e. the ones with a signal-to-noise ratio below a certain level, is typically a very simple and well-known measure for the skilled person to improve the quality of the output sound signals, which can be added to the hearing systems according to claim 1 of the higher-ranking claim requests or its operating methods, respectively, without any difficulties.

4.10 Thus, the method steps of claim 17, being inherent to the hearing system according to claim 1 of the higher-ranking claim requests, do not involve an inventive step. Hence, auxiliary request 3 is not allowable under Article 56 EPC either.

5. *Auxiliary request 4 - inventive step (Article 56 EPC)*

5.1 **Claim 1** of auxiliary request 4 corresponds to claim 1 of auxiliary request 3 and is thus likewise considered to involve an inventive step (see points 4.1 to 4.4 above).

Independent method **claim 17** now includes a method step relating to the user's selection to listen to the output of a single spatial sound signal on the basis of feature (h), thereby effectively limiting the claimed method (see point VIII above).

5.2 The opponent argued in its letter of 14 November 2019 with respect to "auxiliary request 5", which introduced the feature relating to the user selection to listen to the output of a single spatial sound signal according to feature (h), that such selection was already disclosed in document **D8** (referring to paragraph [0028]).

The board, however, holds that paragraph [0028] of D8 refers to different sound inputs not reflecting different *spatial* signals but different *input* channels, i.e. associated with microphones and a wireless receiver, and that the selection of one of these input channels does not refer to a *spatial* signal but to the *source* of the audio signal in general.

- 5.3 In view of the above, auxiliary request 4 is allowable under Article 56 EPC.
- 5.4 Given that no other objections were invoked by the opponent and the board, auxiliary request 4 is considered to comply with all requirements of the EPC.
6. The board further holds that the description as amended before the opposition division does not require further modifications. This was not disputed by the parties.

Order

For these reasons it is decided that:

1. The decision under appeal is set aside.
2. The matter is remitted to the opposition division with the order to maintain the patent in the following version:

Claims 1 to 17 of auxiliary request 4 as filed on 14 November 2019;

Description as adapted before the opposition division:

- Paragraphs 1 to 29, 33 (part), 34 to 85 of the patent specification,
- Paragraphs 30 to 32, 33 (part) filed during the oral proceedings on 14 December 2018;

Drawings: sheets 1 to 4 of the patent specification.

The Registrar:

The Chair:



B. Brückner

K. Bengi-Akyürek

Decision electronically authenticated