Internal distribution code:

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

Datasheet for the decision of 26 November 2021

Case Number: T 2335/17 - 3.4.01

Application Number: 13166981.4

Publication Number: 2772907

IPC: G10L15/00, G06F1/32, H04W52/02, G10L17/00

Language of the proceedings: EN

Title of invention:
Device for activating with voice input

Applicant:
Sony Group Corporation

Headword:
Voice input / Sony

Relevant legal provisions:
EPC Art. 84, 123(2)
RPBA 2020 Art. 13(1), 13(2)

Keyword:
Claims - clarity - auxiliary request (no) - support in the description (no)
Amendments - added subject-matter (yes)
Amendment after summons - exceptional circumstances (yes)
Amendment to appeal case - amendment gives rise to new objections (yes)

Europäisches Patentamt
European Patent Office
Office européen des brevets

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 2335/17 - 3.4.01

D E C I S I O N
of Technical Board of Appeal 3.4.01
of 26 November 2021

Appellant: (Applicant)	Sony Group Corporation 1-7-1 Konan Minato-ku Tokyo 108-0075 (JP)
Representative:	D Young & Co LLP 120 Holborn London EC1N 2DY (GB)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 22 May 2017 refusing European patent application No. 13166981.4 pursuant to Article 97(2) EPC.**

Composition of the Board:

Chair	P. Scriven
Members:	T. Petelski
	C. Almberg

Summary of Facts and Submissions

I. The applicant appealed the Examining Division's decision to refuse the European patent application 13166981.

II. The Examining Division held that the subject-matter of claims 1 and 12 of the main request, and of claims 1 and 9 of the auxiliary request, would have been obvious for the skilled person starting from either of the documents of documents

D2: WO 2010/078386 A1 and
D3: US 2006/074658 A1.

III. In the statement setting out their grounds of appeal, the appellant requested that the decision be set aside and that a patent be granted on the basis of one of the two requests underlying the appealed decision.

IV. In a communication sent with the summons to oral proceedings, the appellant was informed of the Board's provisional opinion. The Board identified problems related to a lack of clarity and support) in both requests. Furthermore, the Board questioned the novelty of claims 1 and 12 of the main request over D3, and agreed with the Examining Division's conclusions that the subject-matter of claims 1 and 12 of the main request, and of claims 1 and 9 of the auxiliary request, lacked an inventive step in view of each of D2 and D3 as starting points.

V. In response, the appellant submitted a new main request. The former main and auxiliary requests were re-filed and re-labelled as the first and second auxiliary requests.

VI. During oral proceedings, which were held by video link, the appellant submitted a new main request to replace the previous one.

VII. This decision is based on

- the main request filed during oral proceedings before the Board,
- the first and second auxiliary requests, re-filed in response to the Board's provisional opinion.

VIII. Claim 1 of the main request reads:

An information processing apparatus comprising:

a microphone (105) configured to detect input voice instructions;

a memory (150) configured to store a plurality of keywords, a portion of the plurality of keywords being associated with a predetermined application of the information processing apparatus;

host processor circuitry (110) and

voice processing circuitry (103) configured to:

match an input voice instruction with a voice of a user by comparing the input voice instruction with the plurality of keywords stored in the memory; and

activate the host processor circuitry (110) when the input voice instruction corresponds to one of the plurality of keywords stored in the memory and is spoken by an authorized user such that the host processor circuitry (110) is in an active state, wherein the host processor circuitry (110) is in a suspended state when the voice processing circuitry (103) performs the match of the input voice instruction with a voice of an authorized user;

and, wherein when the host processor circuitry is activated by the voice processing circuitry (103), the host processor circuitry (110) is configured to:

receive information from the voice processing circuitry (103) regarding a successful match of the input voice instruction and one of the plurality of keywords stored in the memory;

search for an application associated with the one of the plurality of keywords; and

execute the application associated with the one of the plurality of keywords stored in

the memory upon a successful keyword search; and

process further input voice instructions spoken by any user when in the active state by verifying whether a keyword associated with a predetermined application of the information processing apparatus is input to the voice processing circuitry (103).

IX. Claim 1 of the first auxiliary request reads:

An information processing apparatus comprising:

a microphone (105) configured to detect input voice instructions;

a memory (150) configured to store a plurality of keywords, each keyword being associated with a predetermined function of the information processing apparatus;

voice processing circuitry (103) configured to:

compare an input voice instruction with the plurality of keywords stored in the memory; and

activate host processor circuitry (110), which is configured to execute an application based on the comparison, and further wherein the voice processing circuitry (103) is configured to match an

input voice with a registered voice of an authorized user, and an input keyword to a predetermined keyword, and the host processor circuitry (110) is configured to execute the application by any user wherein the host processor circuitry (110) is in a suspended state when the voice processing circuitry (103) performs the comparison.

X. Claim 1 of the second auxiliary request adds to claim 1 of the first auxiliary request the feature

... wherein: the voice instructions are modulated by pulse code modulation before being stored in the memory at a fixed time period and the voice processing circuitry is in an active state for a fraction of the fixed time period.

XI. The features of the respective independent method claims correspond to the features of the independent apparatus claims 1.

XII. The appellant's arguments, in so far as relevant for the decision, are set out in the Reasons, below.

Reasons for the Decision

The application

1. The application is concerned with voice control in (mobile) devices. The goal lies in saving power. This is achieved by suspending a host processor of the mobile device when it is not in use, and activating it only when a voice processor has recognized an activation keyword uttered by a user (e.g. "hello" or "start a call"). In those embodiments covered by the claims (relating to Figures 10 - 13), a security mechanism is created. The host processor is only activated if the activation keyword is spoken by an "authorized user". After activation, there is a time window, during which anyone (whether an "authorized user" or not) can give commands.

Main request - admission

2. The main request was filed during the oral proceedings before the Board. Its admission is governed by Article 13(2) RPBA 2020.

3. The filing was a reaction to objections under Article 84 EPC, raised by the Board at the oral proceedings against the previous main request (which was withdrawn: see point VI, above). This constitutes exceptional circumstances in the sense of Article 13(2) RPBA 2020. Therefore, the Board will apply the criteria applicable under Article 13(1) RPBA 2020.

4. Prima facie, the amendments to claim 1 give rise to a number of new objections under Articles 123(2) and 84 EPC.

5. Firstly, the amended feature *a portion of the plurality of keywords being associated with a predetermined application* is, apparently, not originally disclosed and not supported by the description. The original application and the present description only disclose embodiments in which the remaining portion of keywords is used exclusively for activation, without being related to an application. Such keywords are important where the authorized user merely activates the device, for other users to trigger an application (original application, page 14, lines 4 - 19).

6. According to the appellant, only keywords associated to an application were relevant for the invention. They were not inextricably linked to the activation keywords. Hence, a definition of the remaining portion of keywords was not necessary.

7. This is not persuasive. The application emphasizes the importance of keywords that are only used to activate the host processor (like "hello"). The skilled person would not omit such keywords from the memory.

8. Secondly, the amendment according to which the voice processing circuit is configured to *match an input voice instruction with a voice of a user*, independently of whether the host processor is active or not, is apparently also not originally disclosed and not supported by the description. According to the original application and the present description, it is an important aspect of the invention that matching of voice instructions only happens while the host

processor is suspended, and not when it is already active. In addition, voice instructions are only matched with the voice of an *authorized* user, not with the voice of (any) user.

9. According to the appellant, the claim as a whole taught that, when the host processor was activated, only the content of the voice instruction was identified. By contrast, voice matching was only performed in the suspended state.

10. The argument is not persuasive. The skilled person would not interpret claim 1 differently from its definition. Claim 1 does not define the voice processor as doing something else when the host processor is activated.

11. Thirdly, the last claim feature defines the host processor as configured to *process further input voice instructions spoken by any user when in the active state* without defining an end of the activation. Hence, once activated, the host processor stays activated for an indefinite length of time, during which any user can give instructions. An indefinitely long activation is not originally disclosed und not supported by the description. The only passage, on which the amendment could be based, is the description of the embodiment relating to Figure 10 (see page 12, line 8, to page 13, line 11, of the original application). It teaches that any user may give instructions only during a (limited) security time that starts running with the activation of the host processor (see page 13, lines 6 - 11). Security is a central aspect of this embodiment. Therefore, the skilled person would not have generalized the above teaching to an indefinite time during which any user can give instructions.

12. According to the appellant, the skilled person, with a mind willing to understand, would recognize that a limited time was not relevant to achieving the intended goals of security and power saving. A power saving was already achieved before the activation of the host processor. Further, the limited security time was not linked to any user giving commands. Hence the skilled person would have generalized the teaching from pages 12 and 13.

13. This argument is not persuasive. The security and the power savings are very much linked to a limited security time, after which the host processor needs to be activated anew.

14. Fourthly, the last claim feature, defining that the host processor is configured to verify whether a keyword *is input to the voice processing circuitry*, is apparently unclear and not originally disclosed. On the one hand, it is not clear, how the host processor should know what is input to the voice processing circuitry. On the other hand, this feature is in contradiction with the previous definitions in claim 1 and with the original application. Claim 1 defines, in agreement with the original application, that *voice instructions* are input to the voice processing circuitry. It is only within the voice processing circuitry that keywords are identified by comparing the voice instructions to keywords stored in a memory. Upon a successful match, the host processor receives information on the identified keyword. Hence, the keywords are *output from*, but not *input to* the voice processing circuitry. They are input only to the host processor.

15. According to the appellant, the skilled person, with a mind willing to understand, would know that the voice instruction contained the keyword that was to be recognized by the voice processing circuitry. This understanding was in agreement with the original disclosure.

16. The interpretive capacity of the skilled person must not be overstretched. In combination with the other unclear features and deviations from the description, the skilled person is in doubt as to the intended scope of the claim. She would, therefore, not know how correctly to interpret this contradictory feature.

17. In view of the above objections, the main request is not admitted into the proceedings.

First auxiliary request - clarity and support

18. The host processor circuitry is not defined as part of the information processing apparatus in claim 1. It is not clear, to what extent the features defining the host processor circuitry are meant to restrict the information processing apparatus.

19. Claim 1 defines that each keyword stored in the memory is *associated with a predetermined function*. The claim also defines that the host processor circuitry is configured *to execute an application* based on the comparison of a voice instruction with the keywords in the memory. It is not clear, whether or how the *application* is related to the *function*.

20. The voice processing circuitry is configured to match *an input keyword to a predetermined keyword*.

Considering the different terms used, it is not clear if and to what extent this feature relates to the previously-defined comparison of an *input voice instruction* with the *plurality of keywords stored in the memory*.

21. According to claim 1, the voice processing circuitry compares input voice instructions with keywords stored in a memory. It activates a host processor based on this comparison (for example, if the comparison has a positive result). The host processor then executes an application *by any user*. The activation is not coupled to the result of the voice matching, which is also performed by the voice processing circuitry. The claim does not define any relation of this voice recognition to the activation of the host processor and to the execution of the application. This is not supported by the description.

22. The embodiments relating to Figures 10 - 13 (the only embodiments supporting voice recognition) describe that the host processor is only activated if the voice of the input instructions matches the voice of an authorized user. Only after activation by an authorized user can any user give instructions. And only during a predefined security time. This security aspect is central to these embodiments, but is not reflected by the claim.

23. It follows from the above that claim 1 is not clear and not supported by the description (Article 84 EPC).

24. The same reasons apply to method claim 12, save in respect of the question of whether the host processor circuitry is part of the claimed apparatus.

25. The appellant did not address these objections in detail but argued that a skilled person, reading the claim with a mind willing to understand, would consider the claims to be clear. Claim 1 defined that the host processor was activated only if an authorized used has spoken an activation keyword. Once activated, any user could give instructions. That was the key concept of the invention.

26. The argument is not persuasive. As explained above, the claim defines an activation of the host processor based on a comparison of the input instruction with the keywords stored in the memory. The voice matching has no relation to the activation. Therefore, the key aspect according to the description is not reflected by the claim.

27. For these reasons, the first auxiliary request is not allowable.

Second auxiliary request

28. The feature distinguishing independent claims 1 and 9 from the respective claims of the first auxiliary request does not overcome the objections of lack of clarity and lack of support that apply to that request (Article 84 EPC).

29. The appellant pointed to the arguments that were provided for the first auxiliary request. These arguments are not convincing, for the same reasons.

30. Hence, also the second auxiliary request is not allowable.

Order

For these reasons it is decided that:

The appeal is dismissed.

The Registrar: The Chair:

D. Meyfarth P. Scriven

Decision electronically authenticated