Article

Beyond sense-specific processing: decoding texture in the brain from touch and sonified movement

Texture exploration inside the MRI scanner

Haptic

Audio

Sonified movement
(Audio-Haptic condition)

Multivoxel pattern analysis
To decode texture in the brain

Train ① Test
Within
Haptic ←→ Haptic
Cross ②
Audio ←→ Audio
Bimodal ——→ Bimodal

① Haptic Audio Bimodal
S1 S1 S1
S2 S2 S2
A1 A1 A1

② M1 S1 S2
A1

Within modality decoding
primary sensory cortices contain the signature of non-dominant senses

Cross-modality decoding
shared spatial patterns of activity in M1, A1, S2 and A1

C. Landelle, J. Caron-Guyon, B. Nazarian, ..., O. Félician, J. Danna, A. Kavounoudias

anne.kavounoudias@univ-amu. fr

Highlights
Primary sensory cortices contain signatures of the other non-dominant senses

Auditory and somatosensory cortices host audio-tactile multisensory integration

Primary sensory cortices share representation of texture information across senses

Landelle et al., iScience 26, 107965
October 20, 2023 © 2023
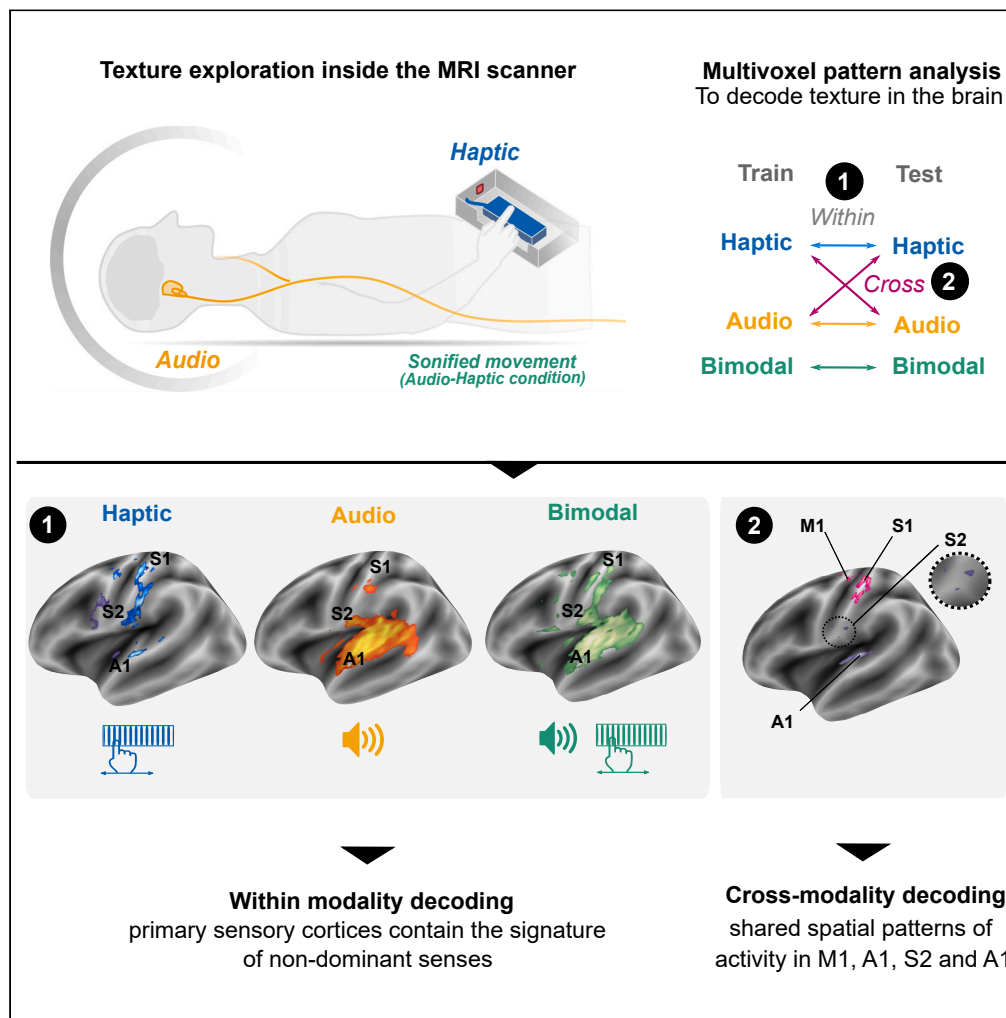https://doi.org/10.1016/
j.isci.2023.107965

Article

Beyond sense-specific processing: decoding texture in the brain from touch and sonified movement

C. Landelle,[1,2] J. Caron-Guyon,[2,3] B. Nazarian,[4] J.L. Anton,[4] J. Sein,[4] L. Pruvost,[5] M. Amberg,[6] F. Giraud,[6] O. Félician,[7] J. Danna,[2,8] and A. Kavounoudias[2,9,*]

SUMMARY

Texture, a fundamental object attribute, is perceived through multisensory information including touch and auditory cues. Coherent perceptions may rely on shared texture representations across different senses in the brain. To test this hypothesis, we delivered haptic textures coupled with a sound synthesizer to generate real-time textural sounds. Participants completed roughness estimation tasks with haptic, auditory, or bimodal cues in an MRI scanner. Somatosensory, auditory, and visual cortices were all activated during haptic and auditory exploration, challenging the traditional view that primary sensory cortices are sense-specific. Furthermore, audio-tactile integration was found in secondary somatosensory (S2) and primary auditory cortices. Multivariate analyses revealed shared spatial activity patterns in primary motor and somatosensory cortices, for discriminating texture across both modalities. This study indicates that primary areas and S2 have a versatile representation of multisensory textures, which has significant implications for how the brain processes multisensory cues to interact more efficiently with our environment.

INTRODUCTION

The perception of texture is traditionally thought to be mediated by touch and vision. The parietal operculum and the posterior insula have been identified as key regions in the processing of haptic textures.[1–7] However, activation in the occipito–temporal cortex, which overlaps with areas activated during pure visual textures perception, has also been frequently reported. This has led to multiple, still unresolved, interpretations.[1,4–6,8] A common explanation is that these posterior visual activations could be mediated by visual imagery.[9] Alternatively, these activations could reflect texture-selective regions within the posterior visual areas that are commonly activated by both the tactile and visual modalities. If this is the case, then one might expect that these amodal regions would generalize their processes across other modalities, such as audition, allowing for the recognition of objects and their features.

The role of hearing in texture perception and the underlying neural process are not well understood. Man et al.[10] used multivariate pattern analysis (MVPA) to reveal that the posterior auditory cortex, the supramarginal gyrus and the secondary somatosensory cortex (S2) contain object-specific representations from both tactile or auditory cues. Therefore, we can hypothesize that the different properties of an object, such as its texture, might be encoded in similar areas for different sensory sources. For instance, when we touch a surface and hear the interaction of our finger on that object, invariant attributes are inherent to both the tactile and auditory modalities. This audio-tactile interaction in roughness perception has been particularly investigated using psychophysical approaches (for review see Di Stefano and Spence[11]). For example, the friction noise created by exploring a surface has shown to also influence our perception of the surface roughness.[12] Interestingly, texture perception can also be modified from amplified natural sounds or supplementary artificial sounds associated with the movement, a strategy called 'movement sonification'.[13] Landelle et al.[14] applied this method and observed that the perception of texture can be modulated by complex textural sounds, but not with neutral sounds (white noise). Another example is the illusion of the hands feeling softer or rougher, called the parchment-skin illusion, which can also occur when the two hands are rubbed together with distorted auditory feedback.[15,16]

[1]McGill University, McConnell Brain Imaging Centre, Department of Neurology and Neurosurgery, Montreal Neurological Institute, Montreal, QC, Canada
[2]Aix-Marseille Université, CNRS, Laboratoire de Neurosciences Cognitives, LNC UMR 7291, Marseille, France
[3]University of Louvain, Institute for Research in Psychology (IPSY) & Institute of Neuroscience (IoNS), Louvain Bionics Center, Crossmodal Perception and Plasticity Laboratory, Louvain-la-Neuve, Belgium
[4]Aix-Marseille Université, CNRS, Centre IRM-INT@CERIMED, Institut de Neurosciences de la Timone, INT UMR 7289, Marseille, France
[5]Aix-Marseille Université, CNRS, Perception, Représentations, Image, Son, Musique, PRISM UMR 7061, Marseille, France
[6]Université Lille, Laboratoire d'Electrotechnique et d'Electronique de Puissance, EA 2697-L2EP, Lille, France
[7]Aix Marseille Université, INSERM, Institut des Neurosciences des Systèmes, INS UMR 1106, Marseille, France
[8]Université de Toulouse, CNRS, Laboratoire Cognition, Langues, Langage, Ergonomie, CLLE UMR5263, Toulouse, France
[9]Lead contact
*Correspondence: anne.kavounoudias@univ-amu.fr
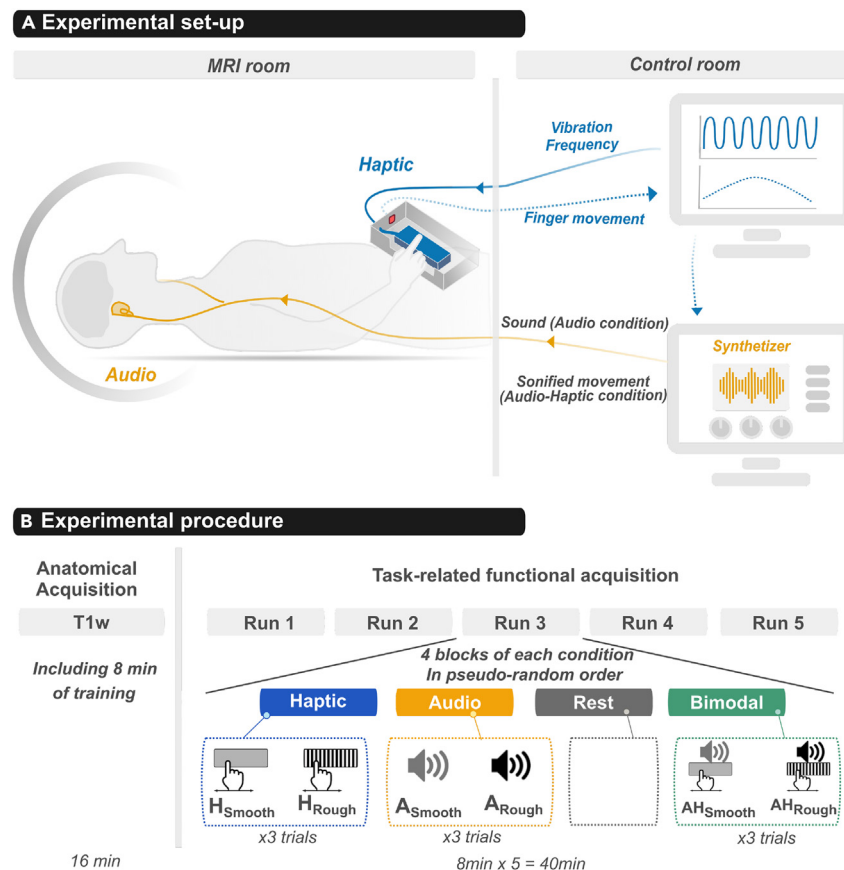https://doi.org/10.1016/j.isci.2023.107965

Figure 1. Illustration of the experimental set-up and procedure

(A) Tactile stimulations (*Haptic*) were delivered by a texture simulator: the StimTac. The participants actively explored the surface of a touchpad delivering a simulated texture while their finger movement was recorded with an optical sensor. Auditory stimulations (*Audio*) were sonified movements generated by a sound synthesizer. Rough or Smooth sounds were produced. Audio-Haptic stimulations (*Bimodal*) were generated by synchronizing the StimTac and the sound synthesizer. In these conditions, the velocity of the participant's finger was sent to the synthesizer to modulate the sounds in real time. A custom program implemented in the NI LabView environment was used to monitor tactile and auditory stimulations in synchronization with MRI acquisitions.

(B) Experimental procedure. The functional acquisitions consisted of 5 runs each including 4 blocks: *Haptic (H)*, *Audio (A)*, *Bimodal (AH)* and *Rest* blocks. The 3 stimulation conditions had two textural levels: Smooth (H_{Smooth}, A_{Smooth}, AH_{Smooth}) or Rough (H_{Rough}, A_{Rough}, AH_{Rough}). Each modality was always presented as a 'block' with 6 consecutive stimuli of the same modality (3x2 textures). Four blocks for each modality were presented in a pseudo-random way including also Rest blocks. In total, a run lasted 8 min.

Although the brain regions involved in audio-tactile texture discrimination have yet to be studied, higher-order brain regions have been found to play a role in integrating information from different modalities in various multisensory tasks.[17,18] In the context of audio-tactile interactions, the posterior superior temporal sulcus (STS) and the fusiform gyrus have been shown to process multisensory information for object-recognition.[19]

Therefore, the current study aimed to uncover and compare the neural regions responsible for pure auditory and pure haptic texture perception, as well as examine the multisensory integration when presented simultaneously. Using MVPA, we also identified regions specifically dedicated to texture coding within each intramodal network. A crossmodal classification approach then allowed us to uncover regions with shared texture representation across modalities.

To achieve this, we developed an innovative method that combined virtual textures in the tactile and auditory domains to elicit different levels of friction. Twenty participants completed a roughness estimation task while inside a 3T scanner, using a magnetic resonance (MR) compatible texture simulation device (StimTac) to deliver haptic textures and a sound synthesizer to generate corresponding auditory textures (Figure 1). This study provided new insights into the neural basis of multisensory texture perception, and how it is modulated by the audio-tactile stimulation.

RESULTS

Roughness perception from audition and touch

During the experiment, participants gave subjective scores of roughness for each of the 2 stimuli tested (Rough, Smooth) in the three conditions (A: Audio, H: Haptic, AH: Audio-Haptic) using a subjective scale ranging from 1 (smoothest texture) to 10 (roughest texture). Seven
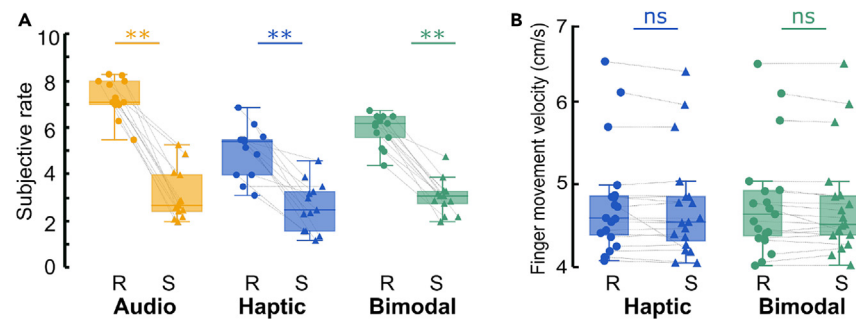
Figure 2. Behavioral measure of texture exploration

(A) Comparison of the subjective roughness rate and (B) finger movement velocity between textures and modalities. Boxplot of the average values (Yellow: Audio, Blue: Haptic, Green: Bimodal) and symbols represent individual values (Rough (R): circle and Smooth (S): triangle). Each box represents the distribution (i.e., from the 25th to the 75th percentile) whereas the medians are represented by the horizontal white line inside the box. Vertical extending lines denote the extreme values within a 1.5 interquartile range. Significant differences were reported, *pcorr <0.0055.

Wilcoxon-paired t-tests were performed to compare the scores within modalities and textures as well as between the bimodal condition and each unimodal condition. Thus, the p values were considered significant for p < 0.007 (Bonferroni correction: 0.05/7). Results revealed significantly higher scores for the Rough compared to Smooth stimuli for all modalities ([A_{rough} vs. A_{smooth}]: 7.10 ± 0.80 vs. 3.12 ± 1.08, v = 91, p = 0.0017; [H_{rough} vs. H_{smooth}]: 5.02 ± 1.08 vs. 2.65 ± 1.04, v = 78, p = 0.0025; [AH_{rough} vs. AH_{smooth}]: 5.92 ± 0.70 vs. 3.05 ± 0.73, v = 91, p = 0.0016, see Figure 2A).

The comparison between the Bimodal and each unimodal condition yielded no significant difference for the Smooth texture [Bimodal vs. Audio: 3.05 ± 0.73 vs. 3.12 ± 1.08, v = 39.5, p = 1; Bimodal vs. Haptic: 3.05 ± 0.73 vs. 2.65 ± 1.04, v = 62, p = 0.020]. When considering the Rough texture, the bimodal condition showed an intermediate perception, being significantly rougher than the Haptic condition [5.92 ± 0.70 vs. 5.02 ± 1.08, v = 87, p = 0.004] and smoother than the Audio condition [5.92 ± 0.70 vs. 7.10 ± 0.80, v = 81, p = 0.002].

Finger movement velocity

Finger movement recordings during fMRI acquisition confirmed that all the participants correctly performed the required task of moving during the Haptic and Bimodal conditions and not moving during the Audio and Rest conditions. Whatever the Haptic or Bimodal condition, the average finger movement velocity of the group was 4.76 ± 0.09 cm/s (mean ± SD; H_{rough}: 4.78 ± 0.10 cm/s, H_{smooth}: 4.73 ± 0.07 cm/s, AH_{rough}: 4.78 ± 0.10 cm/s, AH_{smooth}: 4.75 ± 0.08 cm/s, see Figure 2B). The linear mixed model analysis revealed no significant main effect of roughness (t(4490) = −1.030, p = 0.30) or sensory modality (t(4490) = 0.047, p = 0.96), nor any significant interaction (t(4490) = −0.51, p < 0.61).

Neural correlates of texture perception

Brain activations during haptic, auditory and audio-haptic perception

To identify brain regions that are preferentially recruited during Haptic, Audio or Audio-Haptic perception, we performed a one-way repeated measure ANOVA design ([Haptic>Rest], [Audio>Rest] and [Bimodal>Rest]). As expected, during haptic exploration the most robust effect was observed in the left sensorimotor cortex and the right cerebellum. More precisely, at the cortical level, we found activations within the left primary sensorimotor cortex (SM1, Brodmann area 3 - BA 3), left premotor cortex (PMc, BA 4), bilateral parietal operculum (OP1, OP2) corresponding to the secondary somatosensory cortex (S2) and the right inferior parietal sulcus (IPS, hIP1). The frontal lobe showed activations bilaterally in the supplementary motor area (SMA, BA6), anterior insula and inferior frontal gyrus (IFG, BA44) as well as in the left dorsal PMc (BA6) and right middle frontal gyrus (MFG). The occipital cortex was also found activated within the right fusiform gyrus and primary/secondary visual cortex (V1) as well as the left primary auditory cortices (A1). At the subcortical level, we found activations in the bilateral cerebellum (V, VI, VII, VIII), the bilateral putamen and thalamus as well as in the left colliculus in the brainstem (Figure 3A; Table S1).

During auditory texture perception, the main activation was found bilaterally in A1 (Heschl's gyrus) (Figure 3B; Table S2). Interestingly, most of other activations were common to those found in the haptic condition, including the right IPS, the left PMc and bilateral SMA, IFG, V1, anterior insula, left putamen, and right thalamus. Additional regions showed significant activity in the right hemisphere during the auditory condition and not the haptic condition including right motor regions (PMc and M1) as well as the left IPS. Activation of the cerebellum was less extended than in the Haptic condition and was mainly found in the left cerebellar lobule VIII and in the lobule VII of the vermis. When participants were exposed to an Audio-Haptic stimulation, they recruited extended cortical areas (Figure 3C; Table S3) including the overlapping regions already evidenced by the two unimodal conditions.

Brain areas involved in audio-haptic multisensory integration

To further investigate the brain regions specifically involved in multisensory integration, we performed a conjunction analysis with logical AND. This analysis aimed to reveal brain regions significantly more activated in the bimodal compared to both unimodal conditions
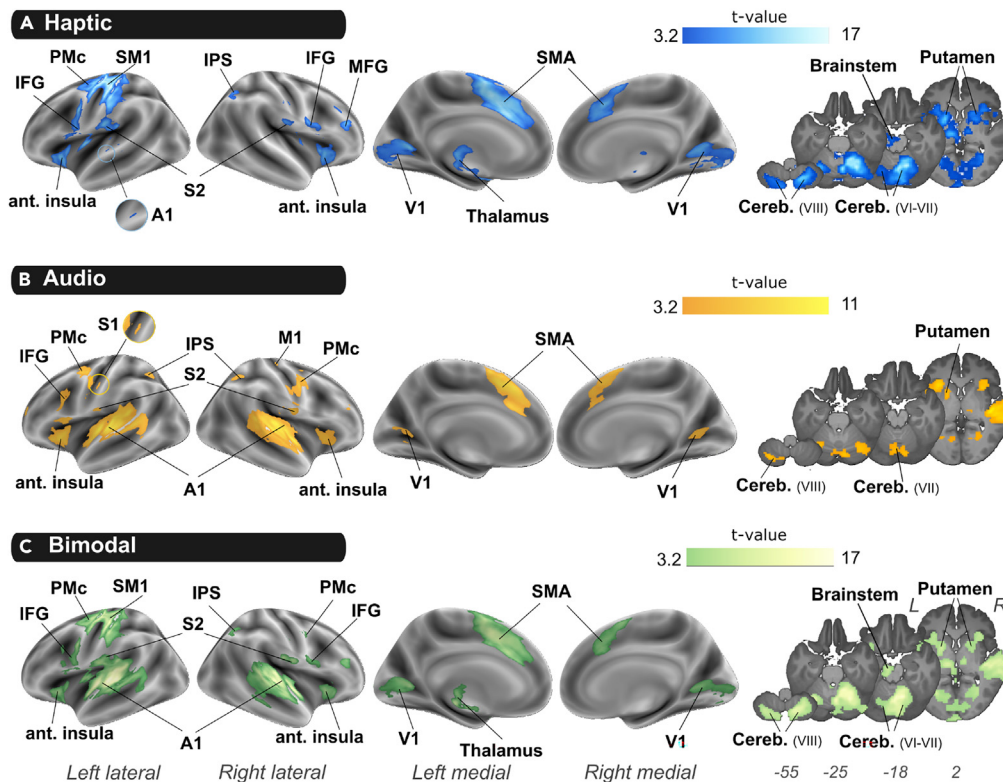
Figure 3. Neural correlates of unimodal texture perception

(A–C) Group statistical t-maps were obtained by contrasting each condition with the rest period: (A) [Haptic > Rest] in Blue, (B) [Audio > Rest] in yellow, (C) [Bimodal > Rest] in green. Statistical maps were thresholded at p < 0.001 uncorrected at the voxel-level and corrected for multiple comparisons at cluster level, FWE: p < 0.05. In the right column, maps in z-plan are overlaid on the average participants' anatomical image (z coordinates in MNI space). L: left hemisphere, R: right, A1: primary auditory cortex, Cereb: cerebellum, IFG: inferior frontal gyrus, IPS: intraparietal sulcus, M1: primary motor cortex, MFG: middle frontal gyrus, SMA: Supplementary motor area, S1: primary somatosensory cortex, S2: secondary somatosensory cortex, SM1: primary sensorimotor area, PMc: premotor cortex, V1: primary visual cortex.

[(AH > A) ∩ (AH > H)]. Activations were found significantly higher in the left A1 and S2 during the bimodal condition compared to the two unimodal conditions (Figure 4; Table S4).

Decoding the texture in the brain

Brain areas decoding texture from auditory and/or haptic modalities

The presence of texture information was evaluated by testing whether the texture classification accuracy was above chance-level (50%) in the A, H and AH conditions separately. In the Haptic condition, group-level non-parametric testing (FWE cluster-corrected) revealed voxels with above-chance classification accuracy in multiple cortical areas, including bilateral S1, S2 and A1, left M1 and right insula (Figure 5A; Table S5) as well as the left PMc and right IPS for a lower statistical threshold (cluster-wise uncorrected p < 0.005). In the Audio condition, texture decoding above-chance was observed in bilateral S1, S2 and A1 (FWE cluster-corrected) as well as the right PMc, the right putamen and the SMA at a lower threshold (cluster-wise uncorrected p < 0.005) (Figure 5A; Table S6). Finally, left S1 and M1 as well as bilateral S2 and A1 displayed significant above-chance classification for textures presented in the bimodal (AH) condition (Figure 5A; Table S7). The right PMc and the SMA also showed significant above-chance classification in bimodal condition at a lower statistical threshold (cluster-wise uncorrected p < 0.005).

Auditory and haptic roughness perception shared neural patterns

We tested whether there was shared texture information across sensory modalities using a crossmodal decoding analysis in which the classifier was trained in one modality and tested in the other. We found significant crossmodal decoding in left S1 and M1 (FWE-cluster corrected p < 0.05). At a lower statistical threshold, a crossmodal decoding was also found in left S2 and A1 (cluster-wise uncorrected p < 0.005) (Figure 5B; Table S8).
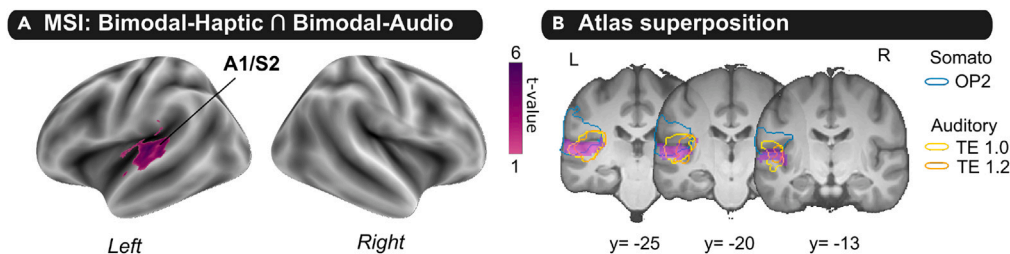
Figure 4. Regions involved in Audio-Haptic multisensory integration

(A) Group statistical t-map was obtained by a global null conjunction analysis between [Bimodal-Haptic] and [Bimodal-Audio] contrasts. t-map was corrected for multiple comparisons (voxel-level uncorrected p < 0.001 and cluster corrected FWE: p < 0.05 FWE).

(B) t-map (in pink) was overlaid on cytoarchitectonic defined maps of the operculum parietale (OP2, in blue) and transverse temporal gyrus of Heschel (TE1.0 in orange and TE1.2, in yellow) respectively described by Eickhoff et al. (2006) and Morosan et al.[20] L: left hemisphere, R: right hemisphere, A1: primary auditory cortex, S2: secondary somatosensory cortex.

DISCUSSION

The present study examined the neural underpinnings of texture perception based on haptic and/or audio information. We found that isolated tactile or auditory texture perception elicited responses not only in their corresponding sensory regions but also in other primary sensory areas. The primary sensory cortices were found to contain signatures of the other non-dominant senses as well as shared representation of texture information across senses. Furthermore, the combination of auditory and tactile information during texture perception revealed that the left A1 and S2 host multisensory processing. Our findings further indicated that motor areas were not only associated with the active haptic exploration of texture, but also involved in the Audio condition, although no movement was being performed.

The primary sensory cortices contain the signature of other senses

The current results showed that haptic and auditory texture perception not only elicited responses to the corresponding primary sensory cortex (S1 or A1, respectively), but also in non-corresponding primary sensory areas (S1 and V1 for the audio condition, A1 and V1 for the haptic condition). This observation corroborates compelling evidence of multimodal activity in primary sensory cortices, traditionally described as sense-specific.[21,22] Numerous electrophysiological recordings in animals have shown that cortical neurons in a given primary sensory area can respond to stimuli from different modalities, such as multisensory neurons in the auditory,[23–27] somatosensory[28,29] and visual[30] cortices. Additionally, there are cortico-cortical and thalamocortical pathways connecting the somatosensory and auditory systems that have been identified using both invasive tracing methods in animals[31–34] and non-invasive structural imaging in humans.[35] Direct projections from A1 to V1 have also been observed in ferrets.[36] These pathways suggest that primary sensory areas are not only hierarchically modulated by higher descending pathways, but that sensory modalities can also influence each other's responses at a lower processing level.

The existence of neural substrates that respond not only to modality-specific stimuli but also to inputs from other modalities may provide an advantage for multisensory processing. For example, activation of these areas in the presence of a single modality could enhance the processing of the corresponding modality presented concomitantly. The present results are in line with this idea, as we found significantly greater activity in the left A1 in the presence of simultaneous audio-tactile stimuli compared to the two unimodal conditions, but not in S1 or V1.

Although the functional significance of such crossmodal responses in primary sensory cortices remains to be determined, one explanation could be the relevance of these areas for texture exploration, regardless of the sensory modality tested. An important common feature during both haptic (vibration under the finger) and auditory (textural sound) texture perception is the frequency content. Several psychophysical experiments evidenced that frequency attributes of the texture are perceptually integrated across touch and audition[15,37–39] suggesting that haptic and auditory frequency features are processed with similar neural mechanisms. This hypothesis is supported by the multivariate analysis, which examined whether there was shared texture information across modalities. Using crossmodal decoding, in which a classifier was trained on data associated with stimulation in one modality and tested in the other modality, we found a significant shared spatial representation of auditory and tactile texture in left M1 and S1, and possibly the left A1 and S2 using a less stringent statistical threshold. This implies that texture is represented at least partially independently of its input modality in these regions. These findings reveal that auditory and tactile texture perception elicited similar spatial patterns of activity not only in their corresponding primary sensory cortex but also in non-corresponding primary sensory cortices.

Finally, the recruitment of V1 in all sensory conditions observed with univariate analysis is consistent with previous fMRI studies investigating the neural bases of texture perception from touch and vision.[1,4,5,40] However, in the present study, this region was not found to support multisensory processing (as indicated by the conjunction analysis) or to contain relevant information for texture exploration (as indicated by the within-modality texture decoding analysis). One possible explanation for the recruitment of V1 herein could be that in daily life, simple surface features of an object (i.e., shape, size …) can be efficiently processed by visual cue whereas more complex properties such as surface roughness are better assessed by tactile information, suggesting that both visual and tactile processes are engaged for surface exploration, but the performance differs across senses (for review see Whitaker et al.[41]). Alternatively, it has been previously proposed that texture
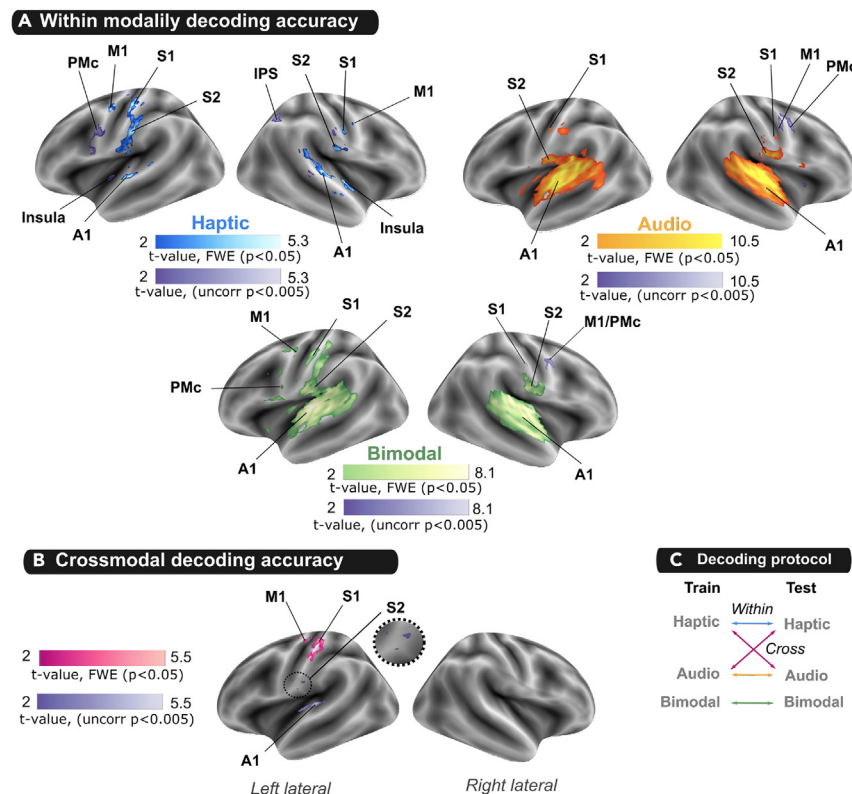
Figure 5. Decoding the texture perception in the brain

(A) Within-modality decoding accuracy: Group statistical t-maps were obtained from non-parametric permutation-based t-test that tested whether the texture classification accuracy was above chance-level (50%) for haptic (blue), audio (yellow) or bimodal (green) decoding.

(B) Crossmodal decoding accuracy: statistical t-maps were obtained from non-parametric permutation-based t-test that tested whether the texture classification accuracy was above chance-level (50%) when the decoder was trained in Audio or Haptic condition and tested in Haptic or Audio condition respectively.

(C) Schematic illustration of the different decoding protocols. Statistical maps were voxel-level uncorrected p < 0.001 and corrected for multiple comparisons at cluster level FWE: p < 0.05 in pink or cluster-wise uncorrected p < 0.005 in purple. A1: primary auditory cortex, IPS: Inferior parietal sulcus, M1: primary motor cortex, PMc: premotor cortex, S1: primary somatosensory cortex, S2: secondary somatosensory cortex.

exploration may also trigger visual imagery as it is common to explore texture with concomitant tactile and visual inputs. However, in contrast to previous studies on tactile shape exploration,[4,9] activation of visual areas did not appear to be specific to the texture since the classifier was not able to discriminate the rough versus smooth conditions within the visual cortex. This could be due to the nature of the stimuli we used, which may not have been easily represented visually. Indeed, Vetter et al.[42] found that the perception of natural sounds like birds singing or a talking crowd can be easily decoded in V1. Future studies using haptic and sound textures that are more visually discriminative should be used to test whether visual information in V1 can be useful for texture decoding in those cases. Although we cannot completely rule out the possibility of visual imagery, participants did not report this strategy during the task.

A1 and S2 host multisensory integration for texture discrimination

A distinctive feature of the present experimental protocol was the integration of motion sonification, aimed at facilitating audio-tactile integration processes. Notably, in the bimodal condition, finger movement sonification generated auditory feedback, creating a compelling perception that the sounds originated from haptic exploration. To rule out the effect of finger movement in multisensory integration measurement, we conducted a conjunction analysis to identify brain regions with greater activation in the bimodal conditions compared to the two unimodal conditions. Given that the finger movement was consistent across the bimodal and haptic conditions, the contrast $[(AH_{Rough}+ AH_{Smooth}) > (H_{Rough}+ H_{Smooth})]]$ helped to eliminate any confounding effects attributable to finger movement in the analysis of the multisensory integration. The conjunction analysis revealed that left A1 and S2 were more activated in the bimodal condition compared to the two unimodal conditions, regardless of the nature of the stimuli (Rough or Smooth). These findings are in line with numerous previous studies (for review see Scheliga et al.[43]), that have demonstrated multisensory integration in the auditory cortex for various types of stimuli such as visuo-tactile,[44] audiovisual,[45,46] audio-tactile[10,23,25] and proprio-tactile[47] stimuli. Additionally, functional connectivity studies have shown that the parietal operculum is a connector hub between auditory, somatosensory and motor areas.[48] Furthermore, the within-modality MVPA analysis applied on each modality as well as in the bimodal condition, revealed a significant texture decoding in these brain regions.

Notably, significant above-chance classification in left A1 and S2 overlapped between the two unimodal and the bimodal conditions. This observation corroborates the crossmodal decoding analysis, which showed significant classification in left A1 and S2 when we used a more permissive statistical threshold (cluster-wise uncorrected p < 0.005). The shared texture representation in left S2 and A1 suggests that they are multisensory areas that contain shared neural patterns representing texture information. However, it is possible that the low significance is due to the fact that in certain parts of these regions, the texture representation in the two sensory modalities may be processed by independent or only slightly overlapping neural populations. In addition, while a potential learned association can't be completely ruled out, the current audio-haptic presentation is unlikely to foster such learning. Indeed, we verified that there were no significant changes in texture estimation across the five runs in either condition, which would have been expected if an implicit association process had occurred.[49]

Auditory-motor networks interaction for texture perception

To optimize the roughness perception, we used active texture exploration in both the Haptic and Audio-Haptic conditions.[50,51] Indeed, texture perception is known to rely on two types of cues: spatial and temporal.[52] The latter is determined by the vibratory frequency of the skin on the surface and is only available when the skin moves across the surface.

A noteworthy finding in the present study was the recruitment of the motor regions in all three experimental conditions (Haptic, Audio and Bimodal), including the M1, PMc, SMA, the cerebellum and subcortical nuclei. While this observation was strongly expected when participants moved their fingers (Haptic and Bimodal conditions), it was more surprising when participants were passively listening to textural sounds although they remained perfectly still (Audio condition). In fact, finger movements during the Audio condition were prevented by training the participants to hold their finger motionless on the StimTac, in a position at the far-right side of the device where movement of the finger was physically blocked by an obstacle (Figure S2).

The interaction between the auditory and motor systems has been well-described in the field of music perception and production (for review see Zatorre et al.[53]) as well as in the rehabilitation of patients with motor disabilities.[54,55] Additionally, neuroimaging studies have reported motor regions recruitment (M1, SM1, PMc, cerebellum) during speech perception[56] and motion sound detection task.[57–59] The activation of the motor regions during the Audio condition is consistent with behavioral findings from a haptic discrimination task, in which participants' motor behavior was modified when their movement was sonified.[14] The present observation is likely supported by the fact that the sound was dynamically modulated by a biological finger movement on the surface (movement sonification). The MVPA analyses also revealed significant above-chance classification of texture in the left and right M1 in the Haptic and Audio conditions respectively. Taken together, these findings suggest that the motor network is strongly associated with both auditory and haptic texture perception, particularly the primary motor cortex, which appears to contain information for texture decoding.

Conclusion

Collectively, the present data support the assumption that primary sensory cortices contain signatures of other senses and host multisensory integration. Multivariate analyses revealed that S1, S2, A1, and M1 contain information about both auditory and haptic texture information. Indeed, similar spatial patterns of activity are elicited within these regions to discriminate texture roughness from auditory and haptic sources, suggesting an efficient interaction between the senses during texture perception. This is further supported by multisensory integration analyses, which showed that S2 and A1 are hubs that may contain multiple sensory representations of texture and host audio-tactile integration. Overall, these findings have important implications for our understanding of the division of sensory processing between primary cortices, traditionally considered as sense-specific. Specifically, our results suggest that regions classically considered to be exclusively somatosensory or auditory may also be sensitive to other senses and even share similar spatial patterns of activity to process texture features.

Limitations of the study

An inherent limitation of our study pertained to the somewhat constrained subjective ratings primarily due to the experiment's time limitations. Only a 10-point rating scale was provided verbally for each stimulus over a total of 60 stimuli. To enhance our understanding of the correlation between subjective ratings and neural activations, a systematic rating after each trial could have provided valuable insights.

Furthermore, future investigations could yield more robust neural representations by expanding the range of stimulus categories beyond the current binary setup. Including an incongruent bimodal condition could also have provided further insights into the neural decoding of the texture congruency between audio-haptic stimuli.

Finally, adding visual stimuli in the experimental protocol could also help to decipher the potential role of V1 in texture discrimination, in the presence and absence of visual information. Such an extension of the protocol would enrich our comprehension of multisensory integration, encompassing a broader spectrum of all three senses.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2023.107965.

AUTHOR CONTRIBUTIONS

Conceptualization - A.K. C.L., and J.D.
Methodology - C.L., J-C-G., B.N., J-L.A., J.S., J.D., and A.K.
Software - Audio stimuli, L.P., J.D., and B.N.
Software - Tactile stimuli, M.A., F.G., and B.N.
Investigation - C.L., J-C-G., and A.K.
Formal Analysis - C.L., J.CG., J-L. A., J.D., and A.K.
Writing – Original Draft, C.L., and A.K.
Writing - C.L., J.C-G., J.D., and A.K.
Review and Editing - all the authors.
Funding Acquisition - J.D. and A.K.
Supervision - A.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

INCLUSION AND DIVERSITY

We support inclusive, diverse and equitable conduct of research.

REFERENCES

1. Eck, J., Kaas, A.L., and Goebel, R. (2013). Crossmodal interactions of haptic and visual texture information in early sensory cortex. Neuroimage 75, 123–135.

2. Kitada, R., Hashimoto, T., Kochiyama, T., Kito, T., Okada, T., Matsumura, M., Lederman, S.J., and Sadato, N. (2005). Tactile estimation of the roughness of gratings yields a graded response in the human brain: an fMRI study. Neuroimage 25, 90–100. https://doi.org/10.1016/j.neuroimage.2004.11.026.

3. Roland, P.E., O'Sullivan, B., and Kawashima, R. (1998). Shape and roughness activate different somatosensory areas in the human brain. Proc. Natl. Acad. Sci. USA 95, 3295–3300.

4. Sathian, K., Lacey, S., Stilla, R., Gibson, G.O., Deshpande, G., Hu, X., Laconte, S., and Glielmi, C. (2011). Dual pathways for haptic and visual perception of spatial and texture information. Neuroimage 57, 462–475.

5. Stilla, R., and Sathian, K. (2008). Selective visuo-haptic processing of shape and texture. Hum. Brain Mapp. 29, 1123–1138.

6. Simões-Franklin, C., Whitaker, T.A., and Newell, F.N. (2011). Active and passive touch differentially activate somatosensory cortex in texture perception. Hum. Brain Mapp. *32*, 1067–1080.

7. Kaas, A.L., van Mier, H., Visser, M., and Goebel, R. (2013). The neural substrate for working memory of tactile surface texture. Hum. Brain Mapp. *34*, 1148–1162.

8. Merabet, L.B., Hamilton, R., Schlaug, G., Swisher, J.D., Kiriakopoulos, E.T., Pitskel, N.B., Kauffman, T., and Pascual-Leone, A. (2008). Rapid and reversible recruitment of early visual cortex for touch. PLoS One *3*, e3046.

9. Lacey, S., Flueckiger, P., Stilla, R., Lava, M., and Sathian, K. (2010). Object familiarity modulates the relationship between visual object imagery and haptic shape perception. Neuroimage *49*, 1977–1990.

10. Man, K., Damasio, A., Meyer, K., and Kaplan, J.T. (2015). Convergent and invariant object representations for sight, sound, and touch. Hum. Brain Mapp. *36*, 3629–3640.

11. Di Stefano, N., and Spence, C. (2022). Roughness perception: A multisensory/crossmodal perspective. Atten. Percept. Psychophys. *84*, 2087–2114.

12. Lederman, S.J., and Klatzky, R.L. (2004). Multisensory Texture Perception. In The Handbook of Multisensory Processes (The MIT Press), pp. 107–122. https://doi.org/10.7551/mitpress/3422.003.0010.

13. Effenberg, A.O. (2005). Movement Sonification: Effects on Perception and Action. IEEE Multimedia *12*, 53–59. https://doi.org/10.1109/mmul.2005.31.

14. Landelle, C., Danna, J., Nazarian, B., Amberg, M., Giraud, F., Pruvost, L., Kronland-Martinet, R., Ystad, S., Aramaki, M., and Kavounoudias, A. (2021). The impact of movement sonification on haptic perception changes with aging. Sci. Rep. *11*, 5124.

15. Guest, S., Catmur, C., Lloyd, D., and Spence, C. (2002). Audiotactile interactions in roughness perception. Exp. Brain Res. *146*, 161–171. https://doi.org/10.1007/s00221-002-1164-z.

16. Jousmäki, V., and Hari, R. (1998). Parchment-skin illusion: sound-biased touch. Curr. Biol. *8*, R190.

17. Kavounoudias, A. (2017). Sensation of movement: A multimodal perception. In Sensation of Movement (Routledge/Taylor & Francis Group), pp. 87–109. https://doi.org/10.4324/9781315627618-6.

18. Klemen, J., and Chambers, C.D. (2012). Current perspectives and methods in studying neural mechanisms of multisensory interactions. Neurosci. Biobehav. Rev. *36*, 111–133.

19. Kassuba, T., Menz, M.M., Röder, B., and Siebner, H.R. (2013). Multisensory interactions between auditory and haptic object recognition. Cereb. Cortex *23*, 1097–1107.

20. Morosan, P., Rademacher, J., Schleicher, A., Amunts, K., Schormann, T., and Zilles, K. (2001). Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage *13*, 684–701.

21. Liang, M., Mouraux, A., Hu, L., and Iannetti, G.D. (2013). Primary sensory cortices contain distinguishable spatial patterns of activity for each sense. Nat. Commun. *4*, 1979. https://doi.org/10.1038/ncomms2979.

22. Ghazanfar, A.A., and Schroeder, C.E. (2006). Is neocortex essentially multisensory? Trends Cogn. Sci. *10*, 278–285.

23. Foxe, J.J., Wylie, G.R., Martinez, A., Schroeder, C.E., Javitt, D.C., Guilfoyle, D., Ritter, W., and Murray, M.M. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. J. Neurophysiol. *88*, 540–543.

24. Fu, K.-M.G., Johnston, T.A., Shah, A.S., Arnold, L., Smiley, J., Hackett, T.A., Garraghty, P.E., and Schroeder, C.E. (2003). Auditory Cortical Neurons Respond to Somatosensory Stimulation. J. Neurosci. *23*, 7510–7515. https://doi.org/10.1523/jneurosci.23-20-07510.2003.

25. Kayser, C., Petkov, C.I., Augath, M., and Logothetis, N.K. (2005). Integration of touch and sound in auditory cortex. Neuron *48*, 373–384.

26. Schroeder, C.E., and Foxe, J.J. (2002). The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Brain Res. Cogn. Brain Res. *14*, 187–198.

27. Schroeder, C.E., Lindsley, R.W., Specht, C., Marcovici, A., Smiley, J.F., and Javitt, D.C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. J. Neurophysiol. *85*, 1322–1327.

28. Kim, S.S., Gomez-Ramirez, M., Thakur, P.H., and Hsiao, S.S. (2015). Multimodal Interactions between Proprioceptive and Cutaneous Signals in Primary Somatosensory Cortex. Neuron *86*, 555–566.

29. Zhou, Y.-D., and Fuster, J.M. (2004). Somatosensory cell response to an auditory cue in a haptic memory task. Behav. Brain Res. *153*, 573–578. https://doi.org/10.1016/j.bbr.2003.12.024.

30. Morrell, F. (1972). Visual system's view of acoustic space. Nature *238*, 44–46.

31. Cappe, C., and Barone, P. (2005). Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. Eur. J. Neurosci. *22*, 2886–2902.

32. Cappe, C., Rouiller, E.M., and Barone, P. (2009). Multisensory anatomical pathways. Hear. Res. *258*, 28–36. https://doi.org/10.1016/j.heares.2009.04.017.

33. Hackett, T.A., De La Mothe, L.A., Ulbert, I., Karmos, G., Smiley, J., and Schroeder, C.E. (2007). Multisensory convergence in auditory cortex, II. Thalamocortical connections of the caudal superior temporal plane. J. Comp. Neurol. *502*, 924–952.

34. Smiley, J.F., Hackett, T.A., Ulbert, I., Karmas, G., Lakatos, P., Javitt, D.C., and Schroeder, C.E. (2007). Multisensory convergence in auditory cortex, I. Cortical connections of the caudal superior temporal plane in macaque monkeys. J. Comp. Neurol. *502*, 894–923.

35. Ro, T., Ellmore, T.M., and Beauchamp, M.S. (2013). A neural link between feeling and hearing. Cereb. Cortex *23*, 1724–1730.

36. Allman, B.L., Bittencourt-Navarrete, R.E., Keniston, L.P., Medina, A.E., Wang, M.Y., and Meredith, M.A. (2008). Do cross-modal projections always result in multisensory integration? Cereb. Cortex *18*, 2066–2076.

37. Bernard, C., Monnoyer, J., Wiertlewski, M., and Ystad, S. (2022). Rhythm perception is shared between audio and haptics. Sci. Rep. *12*, 4188.

38. Crommett, L.E., Pérez-Bellido, A., and Yau, J.M. (2017). Auditory adaptation improves tactile frequency perception. J. Neurophysiol. *117*, 1352–1362.

39. Yau, J.M., Olenczak, J.B., Dammann, J.F., and Bensmaia, S.J. (2009). Temporal frequency channels are linked across audition and touch. Curr. Biol. *19*, 561–566.

40. Kitada, R., Sasaki, A.T., Okamoto, Y., Kochiyama, T., and Sadato, N. (2014). Role of the precuneus in the detection of incongruency between tactile and visual texture information: A functional MRI study. Neuropsychologia *64*, 252–262.

41. Whitaker, T.A., Simões-Franklin, C., Simões-Franklin, C., and Newell, F.N. (2008). Vision and touch: Independent or integrated systems for the perception of texture? Brain Res. *1242*, 59–72. https://doi.org/10.1016/j.brainres.2008.05.037.

42. Vetter, P., Smith, F.W., and Muckli, L. (2014). Decoding sound and imagery content in early visual cortex. Curr. Biol. *24*, 1256–1262.

43. Scheliga, S., Kellermann, T., Lampert, A., Rolke, R., Spehr, M., and Habel, U. (2023). Neural correlates of multisensory integration in the human brain: an ALE meta-analysis. Rev. Neurosci. *34*, 223–245.

44. Macaluso, E., and Driver, J. (2001). Spatial attention and crossmodal interactions between vision and touch. Neuropsychologia *39*, 1304–1316.

45. Macaluso, E., George, N., Dolan, R., Spence, C., and Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. Neuroimage *21*, 725–732.

46. van Atteveldt, N.M., Formisano, E., Blomert, L., and Goebel, R. (2007). The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. Cereb. Cortex *17*, 962–974.

47. Kavounoudias, A., Roll, J.P., Anton, J.L., Nazarian, B., Roth, M., and Roll, R. (2008). Proprio-tactile integration for kinesthetic perception: an fMRI study. Neuropsychologia *46*, 567–575.

48. Sepulcre, J., Sabuncu, M.R., Yeo, T.B., Liu, H., and Johnson, K.A. (2012). Stepwise connectivity of the modal cortex reveals the multimodal organization of the human brain. J. Neurosci. *32*, 10649–10661.

49. Girardi, G., and Nico, D. (2017). Associative cueing of attention through implicit feature-location binding. Acta Psychol. *179*, 54–60.

50. Hollins, M., and Risner, S.R. (2000). Evidence for the duplex theory of tactile texture perception. Percept. Psychophys. *62*, 695–705.

51. Hollins Sliman, J., Bensmaïa Sean, W.M., Bensmaïa, S.J., and Washburn, S. (2009). Vibrotactile adaptation impairs discrimination of fine, but not coarse, textures. Somatosens. Mot. Res. *18*, 253–262. https://doi.org/10.1080/01421590120089640.

52. Katz, D. (1989). The World of Touch (L. Erlbaum).

53. Zatorre, R.J., Chen, J.L., and Penhune, V.B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. Nat. Rev. Neurosci. *8*, 547–558.

54. Danna, J., and Velay, J.-L. (2017). On the Auditory-Proprioception Substitution Hypothesis: Movement Sonification in Two Deafferented Subjects Learning to Write New Characters. Front. Neurosci. *11*, 137.

55. Véron-Delor, L., Pinto, S., Eusebio, A., Azulay, J.-P., Witjas, T., Velay, J.-L., and Danna, J. (2020). Musical sonification improves motor control in Parkinson's disease: a proof of concept with handwriting. Ann. N. Y. Acad. Sci. *1465*, 132–145.

56. Cheung, C., Hamilton, L.S., Johnson, K., and Chang, E.F. (2016). Correction: The auditory representation of speech sounds in human motor cortex. Elife 5, e17181. https://doi.org/10.7554/eLife.17181.

57. Poirier, C., Collignon, O., Scheiber, C., Renier, L., Vanlierde, A., Tranduy, D., Veraart, C., and De Volder, A.G. (2006). Auditory motion perception activates visual motion areas in early blind subjects. Neuroimage 31, 279–285.

58. Reznik, D., Ossmy, O., and Mukamel, R. (2015). Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. J. Neurosci. 35, 2173–2180.

59. Zimmer, U., and Macaluso, E. (2009). Interaural temporal and coherence cues jointly contribute to successful sound movement perception and activation of parietal cortex. Neuroimage 46, 1200–1208.

60. Oldfield, R.C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. Neuropsychologia 9, 97–113. https://doi.org/10.1016/0028-3932(71)90067-4.

61. Aramaki, M., Besson, M., Kronland-Martinet, R., and Ystad, S. (2011). Controlling the Perceived Material in an Impact Sound Synthesizer. IEEE Trans. Audio Speech Lang. Process. 19, 301–314. https://doi.org/10.1109/tasl.2010.2047755.

62. Conan, S., Thoret, E., Aramaki, M., Derrien, O., Gondre, C., Ystad, S., and Kronland-Martinet, R. (2014). An Intuitive Synthesizer of Continuous-Interaction Sounds: Rubbing, Scratching, and Rolling. Comput. Music J. 38, 24–37. https://doi.org/10.1162/comj_a_00266.

63. Feinberg, D.A., Moeller, S., Smith, S.M., Auerbach, E., Ramanna, S., Gunther, M., Glasser, M.F., Miller, K.L., Ugurbil, K., and Yacoub, E. (2010). Multiplexed echo planar imaging for sub-second whole brain FMRI and fast diffusion imaging. PLoS One 5, e15710.

64. Moeller, S., Yacoub, E., Olman, C.A., Auerbach, E., Strupp, J., Harel, N., and Uğurbil, K. (2010). Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. Magn. Reson. Med. 63, 1144–1153.

65. Xu, J., Moeller, S., Auerbach, E.J., Strupp, J., Smith, S.M., Feinberg, D.A., Yacoub, E., and Uğurbil, K. (2013). Evaluation of slice accelerations using multiband echo planar imaging at 3T. Neuroimage 83, 991–1001. https://doi.org/10.1016/j.neuroimage.2013.07.055.

66. Dosenbach, N.U.F., Koller, J.M., Earl, E.A., Miranda-Dominguez, O., Klein, R.L., Van, A.N., Snyder, A.Z., Nagel, B.J., Nigg, J.T., Nguyen, A.L., et al. (2017). Real-time motion analytics during brain MRI improve data quality and reduce costs. Neuroimage 161, 80–93.

67. Gorgolewski, K.J., Auer, T., Calhoun, V.D., Craddock, R.C., Das, S., Duff, E.P., Flandin, G., Ghosh, S.S., Glatard, T., Halchenko, Y.O., et al. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci. Data 3, 160044.

68. Behzadi, Y., Restom, K., Liau, J., and Liu, T.T. (2007). A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. Neuroimage 37, 90–101. https://doi.org/10.1016/j.neuroimage.2007.04.042.

69. Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S., and Turner, R. (1995). Analysis of fMRI time-series revisited. Neuroimage 2, 45–53.

70. Power, J.D., Barnes, K.A., Snyder, A.Z., Schlaggar, B.L., and Petersen, S.E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. Neuroimage 59, 2142–2154.

71. Friston, K.J., Williams, S., Howard, R., Frackowiak, R.S., and Turner, R. (1996). Movement-related effects in fMRI time-series. Magn. Reson. Med. 35, 346–355.

72. Nichols, T., Brett, M., Andersson, J., Wager, T., and Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. Neuroimage 25, 653–660.

73. Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., and Varoquaux, G. (2014). Machine learning for neuroimaging with scikit-learn. Front. Neuroinform. 8, 14.

74. Nichols, T.E., and Holmes, A.P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. Hum. Brain Mapp. 15, 1–25.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Raw data	https://openneuro.org/datasets/ds004743	N/A
Code for MVPA analyses	https://github.com/Centre-IRM-INT/GT-MVPA-nilearn/tree/main/AudioTact_Project	N/A
Software and algorithms		
dcm2niix v1.0.20171215	https://github.com/rordenlab/dcm2niix	RRID:SCR_023517
Brain Imaging Data Structure (BIDS) standard	http://bids.neuroimaging.io	RRID:SCR_016124
FSL	http://www.fmrib.ox.ac.uk/fsl/	RRID:SCR_002823
SPM12	https://www.fil.ion.ucl.ac.uk/spm/	RRID:SCR_007037
GLMflex	http://mrtools.mgh.harvard.edu/index.php/GLM_Flex	N/A
Nilearn toolbox	https://nilearn.github.io/	RRID:SCR_001362
SnPM13	http://nisox.org/Software/SnPM13/	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Anne Kavounoudias (anne.kavounoudias@univ-amu.fr).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- The raw data are available in public repositories hosted at OpenNeuro: https://openneuro.org/datasets/ds004743.
- The code to conduct the MVPA analysis in this study is available at GitHub: https://github.com/Centre-IRM-INT/GT-MVPA-nilearn/tree/main/AudioTact_Project.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Twenty volunteers aged from 20 to 27 years (8 men; 22.6 ± 2.2 years) with no history of neurological or sensorimotor diseases participated in the present study and in a previous related psychophysical study.[14] We verified that all participants were right-handed using Edinburg's manual laterality questionnaire.[60] All participants gave their written consent in accordance with the Helsinki Declaration and the experiment was approved by the Ethics Committee (comité de protection des personnes CPP Ouest II Angers N° 2018-A02607-48). No data regarding ancestry, race, or ethnicity was collected for this study.

METHOD DETAILS

Apparatus and stimuli

The experimental device was the same as the one used in Landelle et al. (2021) and allowed us to deliver texture-like tactile and auditory stimuli. The two types of stimulation were presented in isolation or in combination. Therefore, the participants were tested under three modality conditions: 'Haptic', 'Audio' and 'Bimodal'. Each modality could be presented according to two types of textures: 'Rough' or 'Smooth'.

During Haptic and Bimodal conditions, the participants were asked to move their right finger back and forth on the tactile device (the StimTac, see description in the tactile stimuli section) at around 5 cm/s. Their finger displacements were recorded with an MR-compatible optical sensor at a sampling rate of 200 Hz and filtered with a fourth order low pass Butterworth filter (6 Hz). During the bimodal condition, the instantaneous movement velocity (signal derivative) was extracted in real time and sent to the synthesizer (using the OpenSound protocol, see description in the auditory stimuli section) for modulating the sound accordingly.

For the Audio condition, the participants were instructed not to move their finger. In a previous session, before the scanning, we pre-registered 30 auditory stimuli (15 smooth and 15 rough sounds) modulated by the finger movements of one participant during an actual haptic

exploration of 30 textures simulated by the tactile device (15 smooth and 15 rough textures respectively associated with smooth and rough sounds). Thus, even if the participants did not move, they could hear sounds previously recorded from the sonification of finger movement during haptic explorations. Finally, we verified that the frequency spectrums of the two friction sounds were not masked by the acoustic noise of the EPI sequences used during the functional scanning session (Figure S1).

Tactile stimuli

The 'StimTac' device acts as a texture display by simulated virtual textures. It is animated by a controlled vibration with a few micrometers of amplitude, at an ultrasonic frequency (∼42 kHz), which creates an air gap (squeeze film) that spreads between the user's finger and the whole device's surface. We were able to create a grooved sensation under the participant's finger by alternating two vibration amplitudes of the StimTac at a frequency of 22 Hz. Two kinds of textures were simulated: one very rough texture, qualified as 'Rough' (Alternating amplitudes between 1500 nm and 100 nm) and another slightly striated texture perceived as a smooth surface, qualified as 'Smooth' (Alternating amplitudes between 1500 nm and 1100 nm).

Auditory stimuli

Movement sonification was generated in real time with a synthesizer (Max/MSP, http://cycling74.com). Two synthetized friction sounds were applied from the action-object paradigm.[61,62] This paradigm describes the perceptually relevant audio signal that carries information about the attributes of the action and/or object as they interact. Here, the two sounds were obtained from the evocation of two different actions on the same object (a plastic surface), namely rubbing and squeaking. The differences in the action's attributes affect the perceived roughness, with a squeaking sound evoking greater roughness ('Rough' sound condition) than a rubbing sound ('Smooth' sound condition). Auditory stimuli were delivered through the Sensimetrics S14 MR compatible earphones (https://www.sens.com/products/model-s14/, see Audios S1 and S2).

Synchronizations

The stimulation protocol and finger movement acquisition were implemented using a specific software developed for this study using the NI LabVIEW environment (Figure 1A). The software was synchronized with the MR acquisition using an NI-PXI 6289 multifunction I/O device. A digital input line connected to the TTL (Transistor-Transistor Logic) MR pulses defined the hardware clock source of the stimulation program, whose actions were described in a text file including the commands of both auditory and tactile stimulations. We verified that the StimTac and the optical sensor did not produce artifacts on images or modify the signal-to-noise ratio and vice versa.

Procedure

Participants were familiarized with the device a few weeks before the fMRI experiment in a psychophysical study[14] and with the actual fMRI protocol 1 h before the fMRI experiment. We adjusted the volume of the auditory stimuli according to their individual hearing threshold tested outside the scanner. In addition, we started the MRI acquisitions with an EPI sequence during which each verbal instruction (see description later) and auditory condition (e.g., Rough and Smooth) was presented to the participant to ensure that, despite the noise of the functional MR-sequence, the participants perceived the stimuli and verbal instructions well. None of our participants reported difficulty distinguishing the presented sounds from the MR-scanner sound. Following this test, we started the anatomical acquisitions during which the participants were trained again to perform the task for 8 min. This additional training ensured that under the scanner environment, the participants were still able to perform the task and that the movement of their finger was at a constant speed. Finally, the participants underwent the 5 functional runs of 8 min each. Familiarization before the MRI and during the anatomical acquisition was designed to prevent finger movement errors during the fMRI experiment. Additionally, finger movements were recorded during functional acquisition to ensure that no errors were made by the participants.

Communication between the participant and the experimenter was possible throughout the experiment using MRI compatible microphone (OptoActive) and insert earphones (Sensimetrics).

Experimental MRI protocol

Each run consisted of 4 experimental blocks of conditions (Haptic, Audio, Audio-Haptic, Rest) presented 4 times in a pseudo-random order (Figure 1B). The Rest block consisted in a 14-s rest period during which participants had to do nothing while keeping their eyes closed. The other three consisted of blocks of stimulation during which the participants kept their eyes closed and received 6 consecutive stimuli from the same sensory modality. The Audio (A) blocks consisted of passive listening to a textural sound without moving their finger while during the Haptic (H) and Bimodal (AH) blocks, the participants had to move their right index on the StimTac. In AH blocks, they were simultaneously hearing a congruent textural sound enslaved to their own movement. Before each stimulation block, an auditory instruction was given to the participant. Prior to the H or AH block, the '*Bouge*' (i.e., 'Move') instruction told participants that they should move their finger during each trial until the '*Fixe*' (i.e., 'Immobilize') instruction, which instead told the participant not to move during the A or Rest blocks. Inside A, H and AH blocks, each stimulus was preceded by a beep. This beep allowed the participants to know when they should start moving in H an AH conditions.

The auditory and haptic texture could either be Rough or Smooth (A_{Rough}, H_{Rough}, A_{Smooth}, H_{Smooth}). In the AH block condition, the textural sounds and the haptic textures were always congruent (AH_{Rough} or AH_{Smooth}). A total of 6 consecutive stimuli of the same modality (3x2

textures) of 2.5 s duration followed by an interval inter-stimuli of 2.4 $\pm$ 0.9 s were presented to the participants in each block (total duration about 30s for each block). Over the course of the entire experiment the participants completed 5 runs (about 8 min per run) consisting of 4 blocks per condition x 3 trials x 2 textures and 4 rest periods, for a total of 120 trials per modality. The participants were not informed about the binarity of the stimuli and the experimenter let them believe that many different stimuli would be delivered as in the psychophysical experiment they underwent previously. After each stimulus (i.e., during the interval inter-stimuli), participants were instructed to silently estimate its roughness on a subjective scale ranging from 1 to 10, with 1 corresponding to a very smooth texture and 10 to a very rough texture. They were required to give an oral answer only when the instruction '*Réponse*' (i.e., 'Answer') was auditorily delivered after a trial. A total of 12 responses per run were given corresponding to two responses per stimuli (i.e., A_{Rough}, H_{Rough}, AH_{Rough}, A_{Smooth}, H_{Smooth} or AH_{Smooth}). Therefore, ten estimations per stimulation condition were collected over the entire experience.

In total, the experiment was composed of 5 runs (about 8 min each). Each run was composed of 4 blocks of the 3 experimental conditions lasting 35 s each (including stimuli, interval inter-stimulus and one oral response) and a block of rest period lasting 14 s.

Image acquisition

3-Tesla MRI Scanner (MAGNETOM Prisma, Siemens AG, Erlangen, Germany) with the in-built body-coil for rf excitation and the manufacturer's 64-channel phase-array head coil for reception were used for MRI data acquisition. First, we acquired anatomical images using a high-resolution T1-weighted anatomical image (MPRAGE sequence: repetition time (TR)/inverse time (TI)/echo time (TE) = 2300/925/2.98 ms, voxel size = $1 \times 1 \times 1$ mm^3, slices = 192, fild of view (FOV) = $256 \times 256 \times 192$ mm^3, flip angle = 9°). Blood-oxygen-level-dependent (BOLD) images were acquired with the multiband sequence (Multi-Band EPI C2P v014) provided by the University of Minnesota.[63–65] Whole-brain Fieldmap images were acquired twice with a spin-echo EPI sequence with opposite phase encode directions along the anterior-posterior axis with the following parameters: TR/TE = 7060/59 ms, voxel size = $2.5 \times 2.5 \times 2.5$ mm^3, slices = 56, FOV (Field of View) = $210 \times 210 \times 140$ mm^3, flip angle = 90/180°. BOLD images using a gradient-echo EPI sequence were subsequently acquired with the following parameters: TR = 945 ms, TE = 30 ms, voxel size = $2.5 \times 2.5 \times 2.5$ mm^3, multiband factor = 4, 56 slices, 501 volumes, FOV = $210 \times 210 \times 140$ mm^3, flip angle = 65°. Five runs were acquired using this BOLD protocol.

During the acquisitions head motion was monitored in real-time with FIRMM[66] to improve data quality. In case of significant head motion, the experimenter informed the participant, and the scan could be repeated.

MR image pre-processing

All acquired MR images were converted into NIFTI format using dcm2niix v1.0.20171215 (https://github.com/rordenlab/dcm2niix) and stored following the Brain Imaging Data Structure (BIDS) standard (http://bids.neuroimaging.io).[67] For each participant, FSL (http://fsl.fmrib.ox.ac.uk/fsl/) was used to generate the fieldmap from the two sets of reversed phase encoding SE-EPI using the tool topup. Then all the functional and structural images were pre-processed using SPM12 (Welcome, Department of Imaging Neurosciences, London, UK, http://www.fil.ion.ucl.ac.uk/spm/http://www.fil.ion.ucl.ac.uk/spm/) running on MATLAB 2017 (The Mathworks, Inc., MA, USA).

(1) The functional images were corrected for magnetic field inhomogeneities using fieldmaps (estimated from spin-echo EPI sequence with opposite phase encode directions using FSL *topup*)

(2) fMRI data were then realigned and unwarped to the first image of the first run to correct for head movements between scans. Motion parameters were extracted and visually inspected.

(3) Next, the functional images were coregistered to the T1-weighted (T1w) anatomical image.

(4) T1w structural image was segmented using Cat-12 toolbox into gray matter (GW), white matter (WM) and cerebrospinal fluid (CSF) tissues.

(5) We applied a Component based noise Correction method (CompCor,[68] to eliminate the non-neural aspects of the BOLD signal inside WM and CSF masks. This step consisted in extracting the mean and the 12 first principal components of the BOLD time courses in each functional run (coregistered in T1w space and unsmoothed) for both WM and CSF masks. These components will be used as noise regressors in the first level analysis (see 'MR image analyses')

(6) The Montreal Neurological Institute (MNI) space was employed as a common space. A two-step registration procedure was performed for each subject: i) the T1w image was registered to the MNI template. ii) This warping field was applied to fMRI images in T1w space. This procedure is crucial for inter-subject comparison through a spatial normalization of the individual fMRI data on a common standard space (MNI).

QUANTIFICATION AND STATISTICAL ANALYSIS

Behavioral data analyses

Based on the responses given by participants during the fMRI session we calculated a subjective texture score. Due to technical problems, responses from 14 participants were available and used for the following analysis. The higher the score, the rougher the texture was perceived by the participant. Nine paired Wilcoxon t-tests were assessed, three of them to compare the mean scores between textures (Rough versus Smooth) in each modality (A, H or AH) and six others were used to compare each modality (A versus H, A versus AH and H versus AH) for each kind of texture ('Rough' or 'Smooth'). All tests were corrected for multiple comparisons using Bonferroni correction, with $p_{corrected}$ value (p_{corr}) considered significant for $p < 0.0055$ (0.05/9).

The mean velocity of the back-and-forth displacements of the participants' finger during the haptic exploration of the Stimtac was computed for each H and AH trials (120 trials per modality per participants) for 19 of the 20 participants. A linear mixed model (LMM) analysis was used to compare the mean movement velocity between conditions (AH, H) and texture (Rough, Smooth). This model allowed us to account for the variability of within participant factors (Texture and Modality) and included random intercept for each participant. Main effects and interactions were considered significant for $p < 0.05$.

MR image analyses

Univariate analyses

Pre-processed functional images registered in MNI space and smoothed with an isotropic three-dimensional Gaussian kernel of 5 mm were used for this ananalysis. We defined a design matrix including separate regressors for the 7 conditions of interest (A_{Rough}, A_{Smooth}, H_{Rough}, H_{Smooth}, AH_{Rough}, AH_{Smooth} and Rest) as well as the experimenter's verbal instructions and the participants' answers. The nine regressors were modeled by convolving boxcars functions representing the onset and offset of the stimulation conditions with the canonical Haemodynamic Response Function (HRF) and were included into a General Linear Model (GLM,[69] We estimated head motion by measuring the framewise displacement (FD) which is an indicator that expresses instantaneous head-motion using the three translational and three rotational realignment parameters.[70] We tested the effects of Run on head movements (mean FD $\pm$ sd: Run 1 : 0.13 $\pm$ 0.032; Run 2 : 0.13 $\pm$ 0.036; Run 3 : 0.13 $\pm$ 0.040; Run 4 : 0.13 $\pm$ 0.037; Run 5 : 0.12 $\pm$ 0.034). These values show that nearly all of the volumes had minimal head motion: 12.5% of the volumes having a (in details Run 1 : 12.63%; Run 2 : 11.98%; Run 3 : 12.60%; Run 4 : 13.02%; Run 5 : 12.24%) and less than 1% of the volumes having an FD > 0.5 mm (in details Run 1 : 0.32%; Run 2 : 0.56%; Run 3 : 0.62%; Run 4 : 0.58%; Run 5 : 0.28%).

To further account for the complex problem of the nuisance of head motion on BOLD signal, head movement's parameters estimated by motion realignment procedure were also included into the GLM as nuisance regressors. Specifically, we used 24-parameters autoregressive model of motion[71] including current (i.e., 3 translations and 3 rotations) and past position (6 head motion parameters one time point before) and the 12 corresponding squared items. Finally, the first 12 principal components and the mean signal of each 'nuisance' tissues (CSF and WM) signals obtained during the physiological noise modeling were added in the GLM as regressors of no interest.[68] The time series were high-pass filtered at 128 s.

To examine brain activations related to each modality condition (A, H or AH) for both textures, we created the following first level contrasts for each participant: [(0.5A_{Rough} + 0.5A_{Smooth}) > Rest], [(0.5H_{Rough} + 0.5H_{Smooth}) > Rest], [(0.5AH_{Rough}+0.5AH_{Smooth}) > Rest].

These first-level contrast maps were entered in second-level analyses for each group. The within-groups comparisons were implemented in GLM Flex tool (http://mrtools.mgh.harvard.edu/index.php/GLM_Flex) using a one-way repeated measure ANOVA design including 'Modality' (A, H, AH) as within group factor.

We performed a conjunction analysis to identify candidate multisensory regions that respond more to bimodal stimuli than to each unimodal condition. To do so, we generated the two following first level contrasts: [(AH_{Rough}+ AH_{Smooth}) > (A_{Rough}+A_{Smooth})] and [(AH_{Rough}+ AH_{Smooth}) > (H_{Rough}+ H_{Smooth})]. At the group level, the maps obtained from these two contrasts were entered in a conjunction analysis with logical AND (i.e., the two contrasts should be significant,[72] This analysis explicitly tested which regions are significantly more activated in the bimodal condition than the two unimodal conditions.

Multivariate analyses

The preprocessed functional images projected into T1w native individual space and unsmoothed were used for this analysis. At the first level, a GLM analysis was performed and included separate regressors for each trial (12 trials x 2 textures x 5 runs = 120 trials per Modality) as well as, per run, one regressor for the Rest condition, one for the verbal instruction and one for the participant's answer. As for the univariate GLM, the 24-parameters autoregressive model of head motion and the mean and the 24 first principal components WM and CSF were also included as regressors. The regressors were modeled by convolving boxcars functions with the canonical HRF and high-pass filtered at 128 s. Beta-maps were then estimated for each trial of each modality and texture and used for further MVPA analysis.

Multivariate analyses were processed for each modality (unimodal decoding) and across modalities (crossmodal decoding) using the Nilearn toolbox, a Python package based on scikit-learn library.[73]

For unimodal decoding whole-brain searchlight analyses were performed on single trial beta-maps by centering a spherical neighborhood with a radius of 6 mm at each voxel of the subject-specific whole-brain mask and employing a linear Support Vector Machine (SVM—radial basis function kernel with soft margin parameter C = 1). A leave-two runs out cross-validation scheme was implemented, consisting in training the classifier on three runs and testing it on the other two, resulting in 10 split possibilities. Within-modality decoding were performed to evaluate the brain representation of texture for each condition (A, H, AH). In these cases, we trained the classifier to discriminate the texture (Rough vs. Smooth) on 3 runs and tested it on the 2 remaining runs from the same modality, resulting in 10 cross-validation folds. The same procedure was used for crossmodal decoding except that the decoding was performed on normalized beta-maps (z-scored across runs and modalities). This helps to control for differences in response amplitude to different stimuli across sensory modalities and, thus, ensures that the MVPA classification is linked to differences in the spatial pattern of the activity and not BOLD amplitude differences.[21] The crossmodal decoding tests whether there is shared texture information across sensory modalities (A and H). More specifically, the classifier is trained to discriminate texture (Rough vs. Smooth) on 3 runs in one modality (e.g., A) and tested on the 2 remaining runs in the other modality (e.g., H), and vice versa (i.e., the two directions were investigated). For each participant, each decoding analysis yielded 10 maps representing the classification accuracy per split. The chance level (50%) was subtracted from these maps, which were then averaged across

splits (i.e.,., yielding to one accuracy map per decoding analysis per subject). These final accuracy maps were normalized to the MNI template and smoothed (Gaussian smoothing kernel of 3 mm) before being inserted into a non-parametric permutation-based t-test group analysis (SnPM13.1.06; http://nisox.org/Software/SnPM13/,.[74] Three unimodal (Audio, Haptic and Bimodal) and one crossmodal (two directions) second level non-parametric analyses were run in an explicit mask which is the union maps of the two thresholded univariate maps ([Audio-Rest] ∪ [Haptic-Rest]) to restrict the number of voxels to brain regions activated by at least one modality. In practice, we used 5000 permutations and a significant threshold corrected for multiple comparison at cluster level (p-FWE<0.05).