**Supplementary Information**

**Supplementary Note 1. Linear discriminant analysis and classification**

To perform the formant-based vowel recognition task, the linear discriminant analysis (LDA) method was used to map the 14 formant features to feature space of reduced dimensionality, while maximizing the linear separability of the classes. By normalising these values, we obtain 10 DNPU input voltages *V*1-*V*10. Using LDA, we generate one mapping that is kept constant, so that we stay closer to the reservoir computing approach, where inputs to the DNPU are kept fixed during training. As mentioned in the main text, *V*1-*V*10 are ordered from the most (*V*1) to the least (*V*10) separate generated between the classes. Fig. 4b in the main text shows *V*2 vs *V*1 for all 12 vowels, which are used as inputs to the first (cloned) DNPU. Supplementary Fig. 1 shows similar plots for the other voltages (a: *V*4 vs *V*3, inputs to the second (cloned) DNPU, b: *V*6 vs *V*5, inputs to the third (cloned) DNPU, c: *V*8 vs *V*7, inputs to the fourth (cloned) DNPU, d: *V*10 vs *V*9, inputs to the fifth (cloned) DNPU). The data separability in the output of the LDA layer of *V*i gradually decreases with *i* as can be seen from the overlapping sets of labels.

Chart, scatter chart

Description automatically generated

***Supplementary Fig 1:******a*** *Inputs and corresponding vowel labels for the second (cloned) DNPU.* ***b*** *Same, for the third (cloned) DNPU.* ***c*** *Same, for the fourth (cloned) DNPU.* ***d*** *Same, for the fifth (cloned) DNPU.*

**Supplementary Note 2. DNPU as tuneable activation function**

It is also possible to replace the LDA layer by a trainable linear layer. Using the off-chip gradient descent-based training method, this can be achieved by increasing the size of the combined ANN model and treating the formant features as the input of this combined ANN model, where the DNPU acts as tuneable activation function. In this case L2 regularization is used to keep the input voltages inside the DNPU voltage range (in the rare case that an input voltage falls outside the DNPU voltage range the output is classified as false). This alternative method is applied with the surrogate model (SM) of the DNPU. Supplementary Fig. 2 gives classification accuracies for this alternative method using *N* (cloned) DNPU SMs that can be compared to the classification accuracies in Fig. 4d of the main text, where the LDA is used. For a single DNPU SM (*N* = 1) the classification accuracy for the alternative method is worse than when using the LDA, because in the LDA the best classifying inputs *V*1 and *V*2 are used for this case. However, with increasing *N* the alternative method starts to outperform the corresponding LDA-based method, finally increasing the classification accuracy from 90.6% to 92.6%.

Chart, bar chart

Description automatically generated

***Supplementary Fig 2:*** *Classification accuracies of the vowel recognition task using various numbers of (cloned) DNPU surrogate models when the linear layer that maps formant features to voltages is retrained for each number of DNPUs.*