Loss landscape for cow y-translation
(CLIP model = ViT-L/14)

Legend:
- "a cow sitting under a table"
- "a cow floating above a table"

Y-axis: CLIP loss

"a cow sitting under a table"
(s = -0.30)

"a cow floating above a table"
(s = 0.46)