

Comparing sensory properties of words between English, Dutch, and Italian

Anniaka Tjuka, annika_tjuka@eva.mpg.de, [@AnnikaTjuka](https://twitter.com/AnnikaTjuka)

MAX PLANCK INSTITUTE
FOR EVOLUTIONARY ANTHROPOLOGY

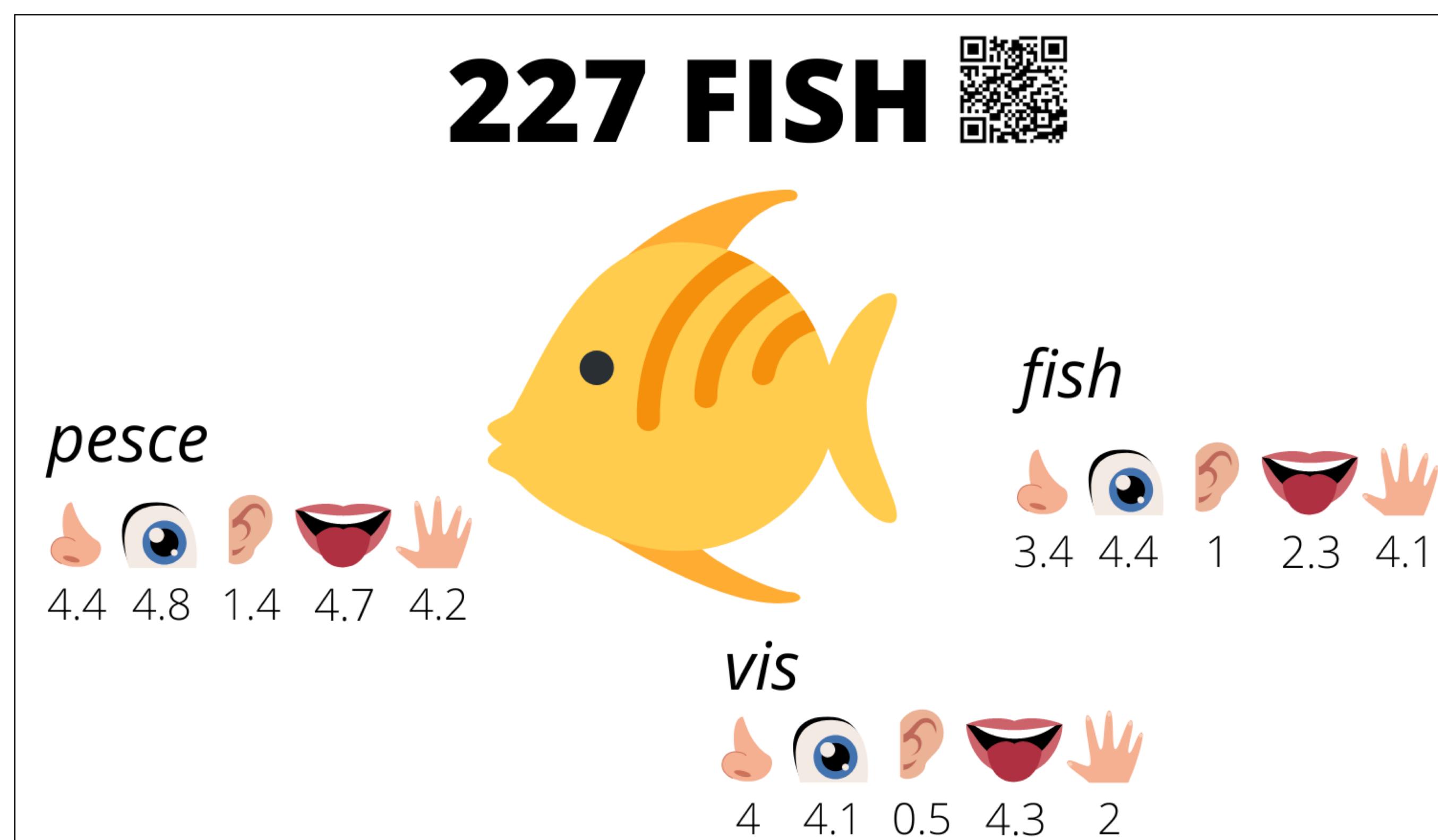


Figure 1. The concept set 227 FISH from Concepticon (<https://concepticon.cld.org/parameters/227>, List et al., 2016) and the words mapped to the concept set. The numbers represent ratings on sensory modality for English (Lynott et al., 2020), Dutch (Speed & Brysbaert, 2022), and Italian (Vergallito et al., 2020).

Introduction

The implications of language diversity for the mental representation and processing of language are of great importance to cognitive science and have received more attention in recent years (Kemmerer, 2019). Variation in word meaning is an inevitable phenomenon that needs further investigation to broaden our understanding of human minds. It is still an open question whether words in different languages that refer to the same concept are represented similarly. Linguists have established the Concepticon database that offers standardized concept sets linked to the respective word in a given language to make judgments about historical relationships between languages (List et al., 2016). Due to expert linguists' informed decisions, this cross-linguistic database has many advantages over automatic translations of words. Cognitive scientists can benefit from a recent extension of the data – the NoRaRe database (Tjuka et al., 2022) – which includes 65 semantic properties offering information on frequency, age of acquisition, and other psycholinguistic measures for 40 languages. The present study used this cross-linguistic database to examine the question of whether the sensory properties of words are similar across English, Dutch, and Italian.

Material and Method

The study compared sensory modality ratings of five dimensions: haptic, visual, olfactory, gustatory, and auditory. The ratings were based on a 5-point scale and were collected for English (Lynott et al., 2020), Dutch (Speed & Brysbaert, 2022), and Italian (Vergallito et al., 2020). The word lists from these studies were mapped to the standardized concept sets in Concepticon (List et al., 2016) to enable cross-linguistic comparison. For example, the words *fish*, *vis*, and *pesce* were mapped to the concept set 227 FISH (see Fig. 1). The mapping was based on a workflow introduced for the NoRaRe database (Tjuka et al., 2022) which maps large numbers of words automatically to the concept sets. The words are linked to the respective ratings for each of the five sensory modalities. The ratings for each language pair are compared on the basis of the concept sets which occur across two word lists (Italian-English: 500 words; Italian-Dutch: 198 words; English-Dutch: 738 words). The words are mainly nouns of the basic vocabulary.

Table 1. Results of the comparison of sensory modality ratings between English, Dutch and Italian based on the data in the NoRaRe database (Tjuka et al., 2022).

Language pair	Words	Sensory modality	R
Italian-English	500 (nouns: 380, verbs: 28, adjectives: 92)	auditory haptic visual gustatory olfactory	0.86 0.85 0.79 0.83 0.83
Italian-Dutch	198 (nouns: 139, verbs: 6, adjectives: 53)	auditory haptic visual gustatory olfactory	0.88 0.83 0.75 0.74 0.78
English-Dutch	738 (nouns: 367, verbs: 28, adjectives: 183, other: 160)	auditory haptic visual gustatory olfactory	0.84 0.77 0.73 0.9 0.83

References

- Kemmerer, D. (2019). Concepts in the brain: The view from cross-linguistic diversity. Oxford, UK: Oxford University Press.
List, J.-M., Cysouw, M., & Forkel, R. (2016). Concepticon: A resource for the linking of concept lists. In N. Calzolari et al. (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation*. Portorož, Slovenia: European Language Resources Association.
Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (2020). The Lancaster Sensorimotor Norms: Multidimensional measures of perceptual and action strength for 40,000 English words. *Behavior Research Methods*, 52.
Speed, L. J., & Brysbaert, M. (2022). Dutch sensory modality norms. *Behavior Research Methods*, 54.
Tjuka, A., Forkel, R., & List, J.-M. (2022). Linking norms, ratings, and relations of words and concepts across multiple language varieties. *Behavior Research Methods*, 54.
Vergallito, A., Petilli, M. A., & Marelli, M. (2020). Perceptual modality norms for 1,121 Italian words: A comparison with concreteness and imageability scores and an analysis of their impact in word processing tasks. *Behavior Research Methods*, 52(4).

Results

The Pearson coefficients for the five sensory modalities were above $R=0.7$ in all language pairs as shown in Table 1 and Figures 2-4. This suggests that the sensory properties of the words are perceived similarly by English, Dutch, and Italian speakers. Interestingly, subtle differences in the individual sensory modalities became apparent. For example, the correlations in the visual modality were the lowest (about $R=0.7$), whereas the Pearson coefficients for the auditory modality were above $R=0.84$ across all languages. The ratings in the gustatory modality were very similar across Dutch and English, but both differed from Italian.

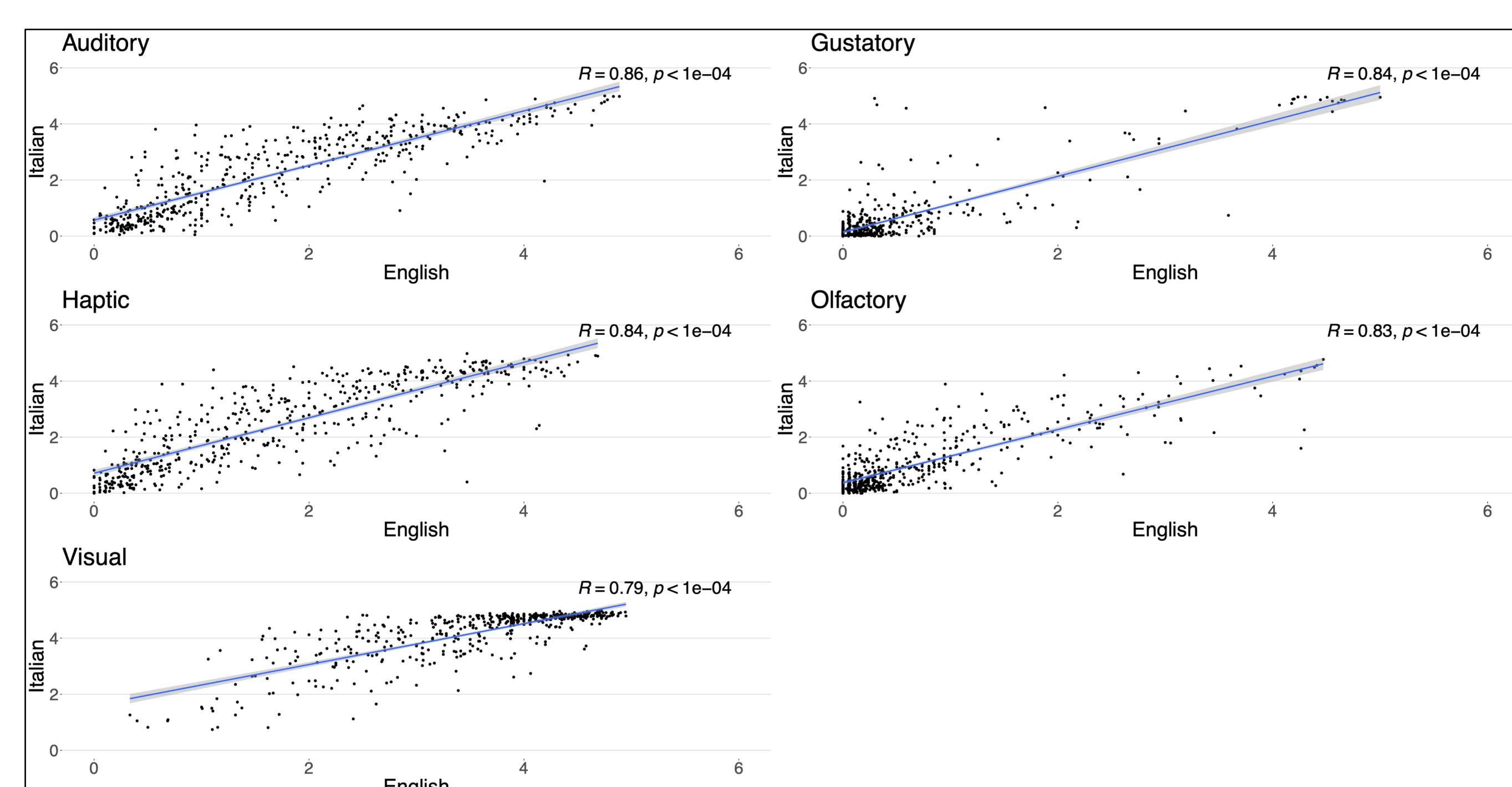


Figure 2. Comparison of sensory modality ratings between Italian and English.

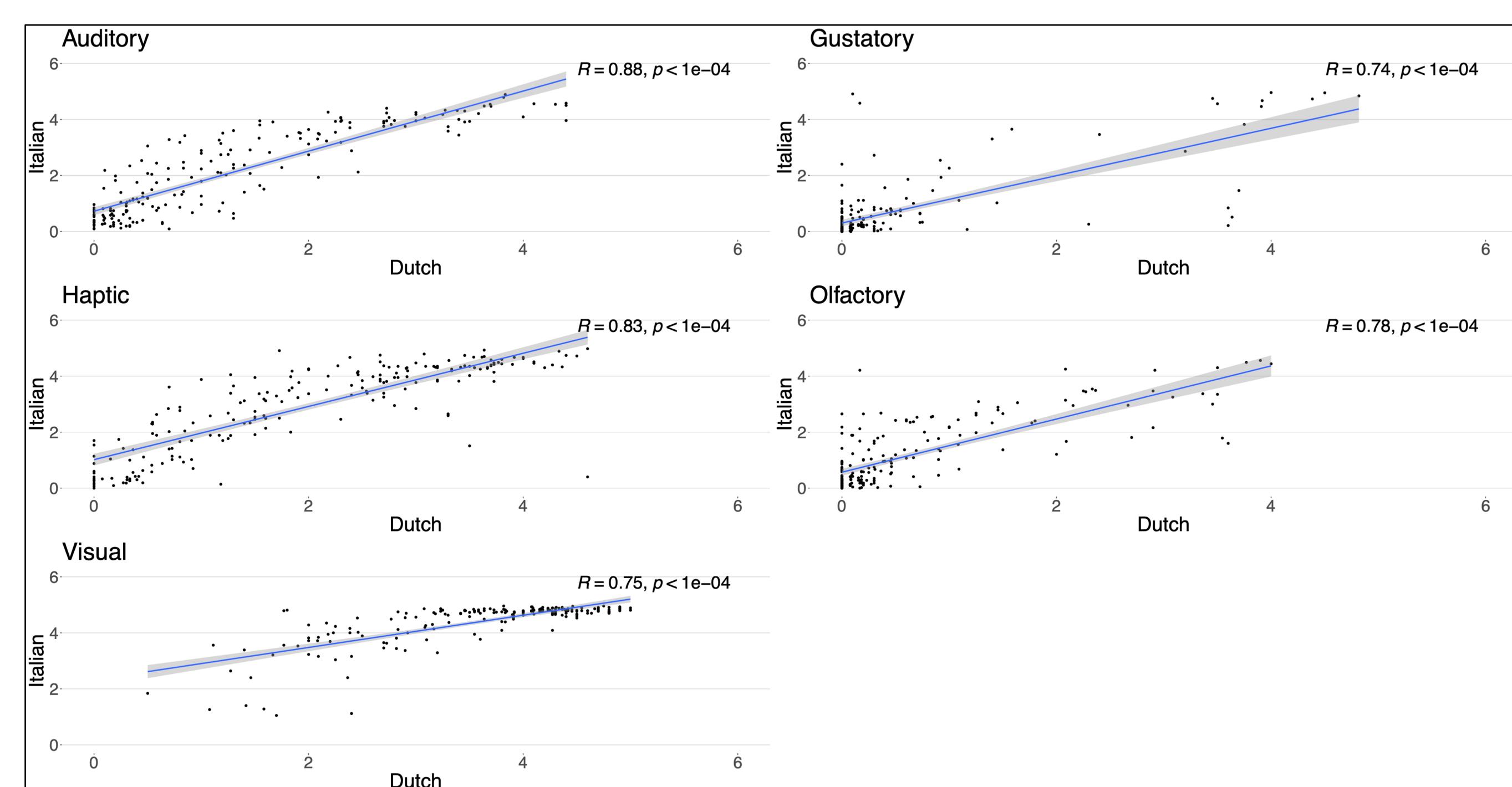


Figure 3. Comparison of sensory modality ratings between Italian and Dutch.

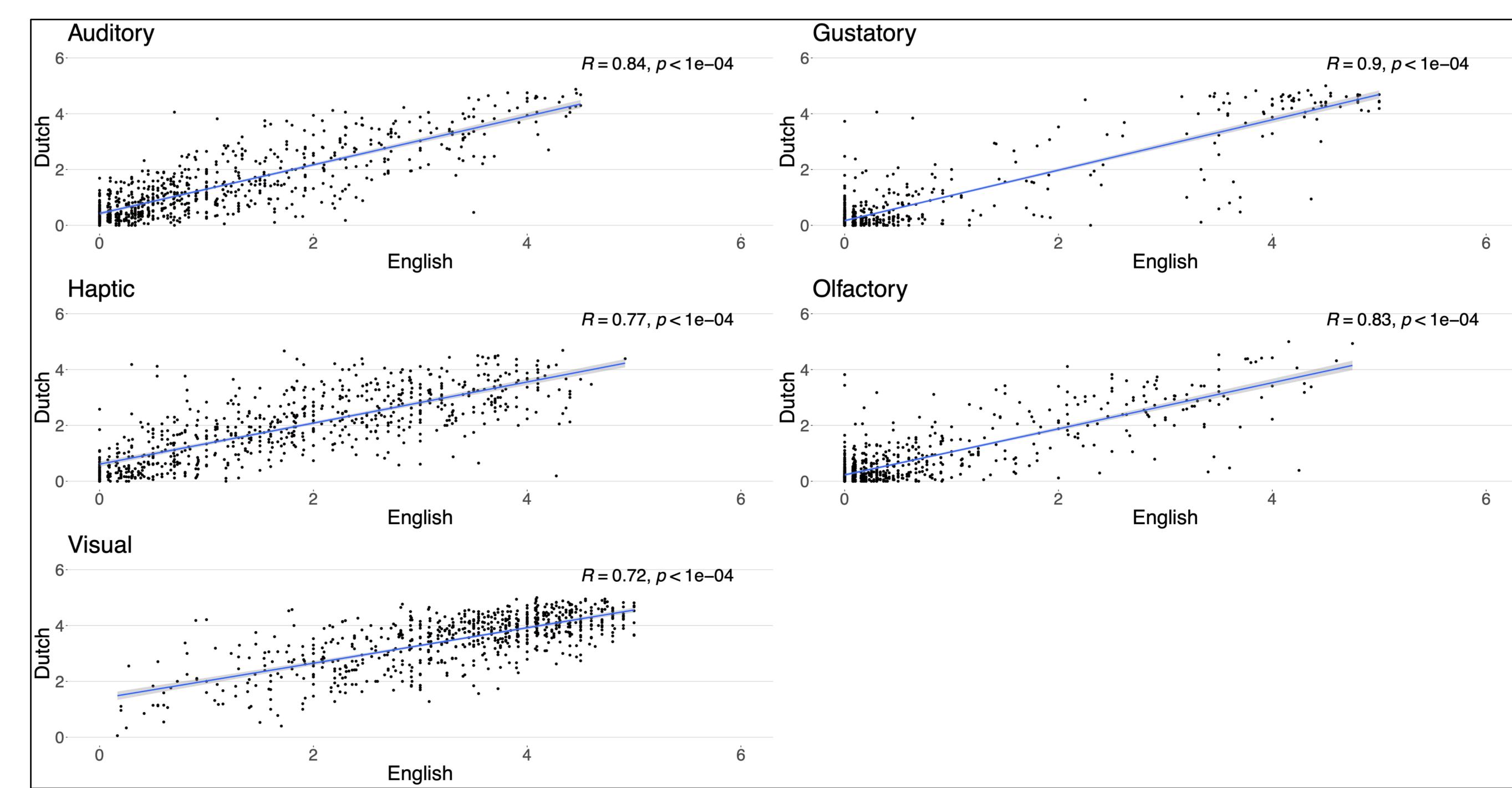


Figure 4. Comparison of sensory modality ratings between English and Dutch.

Implications

The present study focused on a comparison of languages that are closely related and showed that sensory properties of words are perceived similarly across English, Dutch, and Italian speakers. However, additional data for various languages with the same rating scale need to be collected before a general claim can be made about the perception of sensory properties of words across cultures. The implications of such a large-scale study would be far-reaching because it could reveal important differences and similarities in the representation of word meanings across human minds. With the help of cross-linguistic databases, linguists and cognitive scientists can work together to answer big picture questions and generate cross-disciplinary insights about the relation between language, cognition, and culture.