The semantics of spatial demonstratives

Mikkel Wallentin (mikkel@cc.au.dk)
Department of Linguistics, Cognitive Science and Semiotics, Aarhus University, Aarhus, Denmark

Roberta Rocca (roberta.rocca@cc.au.dk)
Department of Linguistics, Cognitive Science and Semiotics, Aarhus University, Aarhus, Denmark

Abstract

Spatial demonstratives (words like *this* and *that*) are thought to primarily be used for carving up space into a peripersonal and extrapersonal domain. However, when given a noun out of context and asked to couple it with a demonstrative, speakers tend to use *this* for manipulable objects (small, harmless, inanimate), while non-manipulable objects (large, harmful, animate) are more likely to be coupled with *that*. Here, we extend these findings and map demonstrative use along a wide spectrum of semantic features. We conducted a large-scale (N = 2197) experiment eliciting demonstratives for 506 words, rated across 65+11 perceptually and cognitively relevant semantic dimensions. We replicated the findings that demonstrative choice is influenced by object manipulability. Demonstrative choice was additionally found to be related to a set of semantic factors, including valence, arousal, loudness, motion, time and more generally, the self. Importantly, demonstrative choices were highly structured across participants, as shown by a strong correlation detected in a split-sample comparison of by-word demonstrative distribution.

Keywords: language, semantics, spatial demonstratives

Introduction

One of the central ways in which language coordinates attention is via spatial demonstratives. Words like the pronouns *this* or *that*, or the adverbs *here* and *there* are among the few undisputed language universals (Diessel, 1999, 2014). They are developmental (Capirci, et al., 1996) and evolutionary (Pagel, et al., 2013) cornerstones of language, and they are among the most frequent words in the lexicon (Leech, et al., 2014).

Demonstratives are *deictic* expressions which in principle can be used to indicate *any* object, and their meaning depends on the context of utterance (Diessel, 1999). During speech monitoring, demonstratives elicit activation in the brain's parietal lobe, suggesting a link to the where/how processing pathways and attention (Rocca, et al., in press). Demonstratives thus form a vital link between language and non-linguistic perception and cognition.

We most often use the proximal demonstrative (*this)* to refer to objects within manual reach and the distal (*that*) for objects beyond (Coventry, et al., 2008). Previous experiments, however, have shown how elicitation of demonstratives not only reveals information about the spatial location of the object, but also relates to the speaker's relationship to the referenced objects (Coventry, et al., 2014; Rocca, et al., 2019a) and to the conversational situation (Peeters & Özyürek, 2016; Rocca, et al., 2019b).

In a recent study (Rocca et al., 2019), Rocca and collaborators introduced the *Demonstrative Choice Task* (*DCT*), a new experimental paradigm where participants are asked to match nouns (e.g. *tiger*, or *apple*) with a demonstrative (i.e. *this* or *that*) without any further context. Across three languages, the authors found that participants consistently use the distal demonstrative (*that*) for a word like *tiger*, whereas they consistently choose *this* for a word like *apple*. This effect was found to be related to the inferred manipulability of the object, both related to inferred perceptual (size) and psychological (harmfulness) semantic dimensions. This is in line with research suggesting that demonstratives are interconnected with kinematic planning (Bonfiglioli, et al., 2009; Caldano & Coventry, 2019; Rocca, et al., 2018) and interactional affordances (Rocca, et al., 2019b), rather than being mere distance indicators.

In this experiment we expanded on this line of research, further validating the DCT and exploring how semantic dimensions other than manipulability affect how speakers choose to couple demonstratives and content words in the absence of context. First, we attempted to establish whether the pattern of demonstrative choices for particular words are reproducible across a large set of words. Secondly, we aimed to replicate Rocca et al.'s previous finding that manipulability affects demonstrative choice. Lastly, we tested if additional semantic domains have an influence on demonstrative choice, thus providing a comprehensive characterization of the relationship between semantics and demonstrative use.

Demonstrative use depends on the establishment of an "origin", serving as the centre of the frame of reference from which an utterance is constructed (Bühler, 1934/2011). The semantic interpretation of *here* and *this* etc. thus presupposes a coordinate system anchored by some entity, usually the speaker's body. However, we also know that spatial demonstratives can be used to denote nonspatial semantic features, such as time (e.g. *this time*), events (*this event*), emotions (*this emotion*), phenomenology (*this experience*) and abstract notions (*this abstraction*), that have no clear spatial anchoring. More generally, as noted by Bühler (1934/2011), deictic reference can be used in an imagination-oriented fashion ("deixis am Phantasma"), i.e. to refer to non-spatial entities such as discourse elements, memories, imagined scenes or other products of "constructive phantasy".

We hypothesize that spatial demonstratives, in the broadest sense, by default will have a strong imagination-oriented function and map onto a coordinate system, not anchored by the physical body, but by the "self" of the speaker. The self includes the speaker's body, but extends it with multiple semantic dimensions, including temporality (i.e. discourse markers such as anaphora), emotions, phenomenology, and social embeddedness (see Hanks, 2009). Our hypothesis is

thus that such dimensions should be observable using the simple Demonstrative Choice Task.

Methods

We conducted a large-scale experiment based on the DCT on the website http://prolific.ac. 2197 native English-speakers participated (Gender: 1364 female, 819 male, 13 other; age: 801 were 18-30 years, 693 were 30-40 years, 347 were 40-50 years, 244 were 50-60 years, and 111 were 60+ years). The study was approved by the Institutional Review Board at Aarhus University.

The study took on average 4 minutes to complete, and participants were rewarded with 0.42 GBP. In the study, participants were presented with 48 or 49 words, selected from a database of 535 words, which have been rated on 65 different semantic dimensions based on neurobiological considerations, comprising sensory, motor, spatial, temporal, affective, social, and cognitive experiences (Binder, et al., 2016). The 535 words were divided into 11 subsets of either 48 or 49 words, and participants were subjected to one subset in a pseudorandomized manner. Similar to Rocca and colleagues' experiment (2019a), participants were asked to couple each word with either the spatial demonstrative *this* or *that* without further context. They were instructed to simply follow their intuition and choose the combination of demonstrative and word they thought fitted best.

Data analysis

Data was analysed in RStudio version 1.1.383 (RStudio Team, 2016). All data and scripts for analyses are available from Open Science Framework: https://osf.io/tqejb/

The 65 semantic dimensions that words are rated along in the Binder dataset can be seen from figures 1 and 2, ordered according to semantic factors. The database is available here: http://www.neuro.mcw.edu/representations/index.html, and the rationale for the choice of these exact features is extensively described in Binder et al. (2016).

One of the aims of the present work was to test the replicability of results from Rocca et al. (2019), where manipulability is argued to play a role in demonstrative choice. The Binder et al. (2016) dataset does not provide an explicit manipulability dimension. We initially attempted to extract a proxy for manipulability by applying principal component analysis and factor analysis on the Binder dimensions. This, however, did not yield a component that could be straightforwardly interpreted as such. We therefore added to our feature set the Lancaster Sensorimotor Norms (available here: https://osf.io/7emr6/). This dataset provides ratings along 11 sensorimotor features for a large body of words (Lynott, et al., In press). The 11 dimensions can also be seen from figure 1 and 2, ordered according to semantic factors (the affix _Lan is appended to differentiate them from features from the Binder dataset).

The overlap between the two databases included 506 out of the original 535 words. All subsequent analyses are conducted on this subset of the data, using the 65+11 semantic features. All feature ratings were standardized to make them comparable. Two features contained missing ratings for particular words. These were imputed using the mean of all other words along that feature.

Factor analysis

To determine the number of latent factors, we used Horn's parallel method (Horn, 1965), implemented in the *psych* package in R. This method compares the scree of factors of the observed data with that of a random matrix and random samples (randomized across rows) of the original data and subtracts out the components that explain less variance than a comparable factor based on non-informative data (see analysis script: https://osf.io/7emr6/ for an illustration). The estimated remaining number of factors using this procedure was 12.

Factor analysis was conducted using Ordinary Least Squares (OLS) to find the minimum residual (minres) solution. Orthogonal rotation (varimax) was applied. The cumulative proportion of variance of the semantic features explained by the 12 factors was 0.75.

Factor labels were given by the authors on the basis of inspection of the features yielding the highest factor loadings (see figures 1 and 2). The 12 factors and the proportion of the variance they explained in the semantic features were: *Vision* (0.14), *Valence* (0.11), *Loudness* (0.09), *Human* (0.06), *Taste/Smell* (0.06), *Motion* (0.06), *Manipulability* (0.06), *Scene* (0.05), *Time* (0.03), *Torso/Legs* (0.03), *Arousal* (0.03), *Self* (0.03) (See figures 1 and 2). It is important to note that these factors and the relative variance they explain do not reflect the general distribution in language or semantics, but only the particular word and feature sample present in the combined databases. The ordering of the factors may therefore be partly specific to these stimuli and features.

The 12 factors were used as predictors in an aggregate-level linear regression analysis investigating the role of semantic dimensions in the distribution of demonstrative choices for words (see below for details).

Aggregate level analyses

At the aggregate level, we used the proportion of proximal demonstratives chosen for each word as outcome variable.

The first aim of the analysis was to investigate the consistency in demonstrative choices across participants and words. We divided the data into two parts and calculated the proportion of proximal demonstratives chosen for each word in both samples. This yielded a vector of 506 proportion values (one per word) per sample. If participants' choices of demonstrative forms for each word were random or highly inconsistent, we would expect the two vectors to be uncorrelated or only very weakly correlated. A strong correlation would speak in favor of participants' coupling of demonstratives and stimulus words being structured and thus, at least to some extent, predictable.

Secondly, we conducted a linear regression analysis with the overall proportion of proximal demonstratives chosen for each word as dependent variable and the 12 factors as independent variables to determine which (if any) semantic factors could be used to predict demonstrative choices.

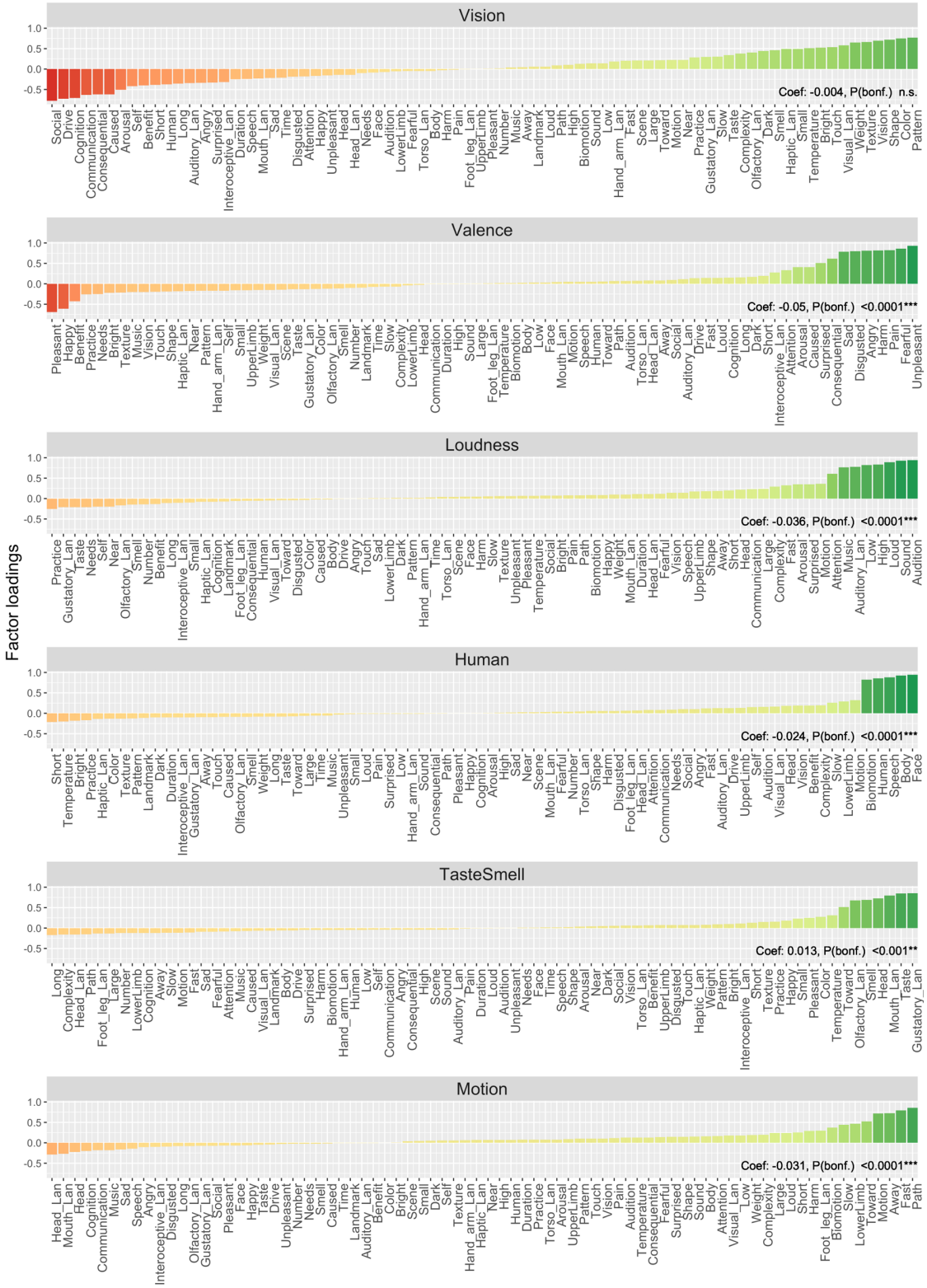

Figure 1. Factor analysis on a combination of Binder and Lancaster features resulted in 12 factors. Here, factors 1-6 are displayed (see figure 2 for factors 7-12), with features ordered by loading. Factors are labelled by the authors. Coefficients reflect regression results. A significant positive coefficient means that positive (green) sematic features are likely to elicit a proximal demonstrative, whereas features with negative (red) loadings tend to elicit distal demonstratives. When the coefficient is negative, the effect of the factor is reversed in the regression, i.e. features with positive loadings (green) are more likely to elicit distal demonstratives.

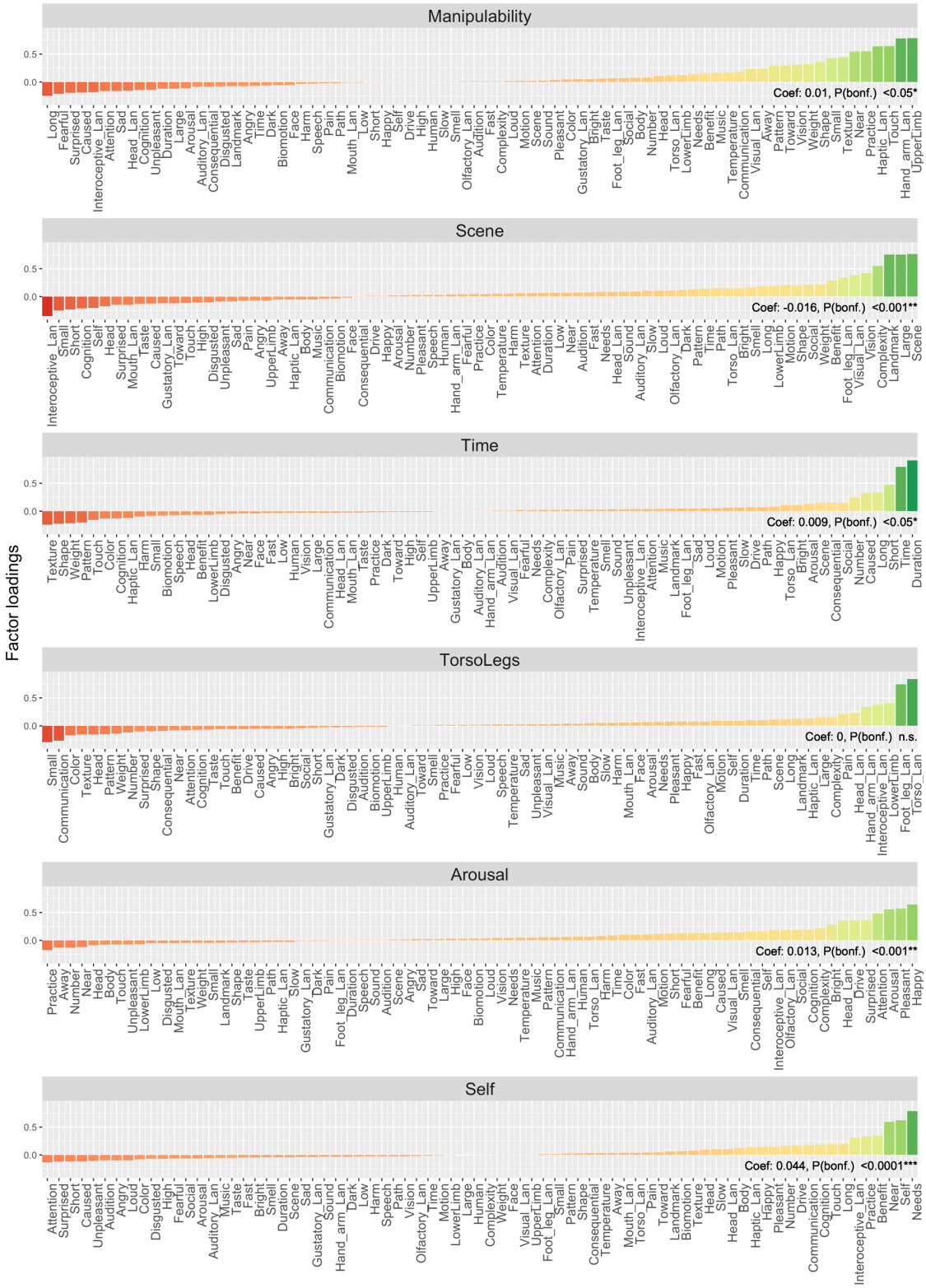

Figure 2. Factor analysis on a combination of Binder and Lancaster features resulted in 12 factors. Here, factors 7-12 are displayed, with features ordered by loading. Factor 7 (top panel) represents *manipulability*, which was hypothesized to explain demonstrative choice. See figure 1 for additional details.

Results

Descriptive results

The overall proportion of proximal/distal demonstratives in the data was 0.465/0.535 (standard deviation across words: 0.114).

Aggregate level analyses

The proportion of proximal demonstratives was highly correlated across the two samples (r=0.82, t(503)=32.7, p<0.0001, figure 3), which speaks in favour of participants' choices of demonstrative forms not being random.

The linear regression model with semantic factors (figures 1-2) as independent variables and overall proportion of proximal demonstratives as dependent variable was highly significant (Adjusted R-squared: 0.6018), indicating that the semantic factors did explain variability in the distribution of proximal and distal demonstratives.

Out of the 12 semantic factors, 10 significantly contributed to the model (p<0.05, Bonferroni corrected): *Valence* (t(493)=-15.6, p<0.0001), *Loudness* (t(493)=-11.3, p<0.0001), *Human* (t(493)=-7.4, p<0.0001), *Taste/Smell* (t(493)=4.0, p<0.001), *Motion* (t(493)=-9.4, p<0.0001), *Manipulability* (t(493)=3.1, p<0.05), *Scene* (t(493)=-4.5, p<0.0001), *Time* (t(493)=2.9, p<0.05), *Arousal* (t(493)=-4.0, p<0.001), and *Self* (t(493)=13.0, p<0.0001). The factors *Vision* and *Torso/Legs* were non-significant (p>0.05). Positive regression coefficients (see figure 1 & 2) and t-values indicate that the factor contributes positively to the choice of proximal demonstratives (i.e. elicits *this* more often), whereas negative regression coefficients and t-values indicate that the factor contributes negatively to the choice of proximal demonstratives (i.e. elicits *that* more often).

A linear combination of factor loadings and regression coefficients for the 10 significant components allows us to project the effects back into feature space (figure 4). This shows how positive valence is an important driver of demonstrative choice, in combination with self-relatedness, proximity and features relevant for manipulability. Negative valence, motion and loudness drive choices towards the distal demonstrative.

Discussion

In this study, we documented that a seemingly meaningless task, such as pairing either a proximal or distal demonstrative with a content word, yields highly reproducible results. The proportion of proximal demonstratives for specific words in one data split closely matched the proportion in the complementary participant sample (see figure 3).

We also replicated the result (Rocca, et al., 2019a) that affordances for haptic interaction (manipulability) predict the choice of proximal demonstratives. This effect, however, was found in a combination with nine additional semantic factors that also significantly contributed to the choice.

On the face of it, the manipulability effect in the current experiment seems less pronounced than the one found in our previous study (Rocca, et al., 2019a). The regression coefficient is smaller than several other factors (see figures 1-2), suggesting that the manipulability factor is not the main

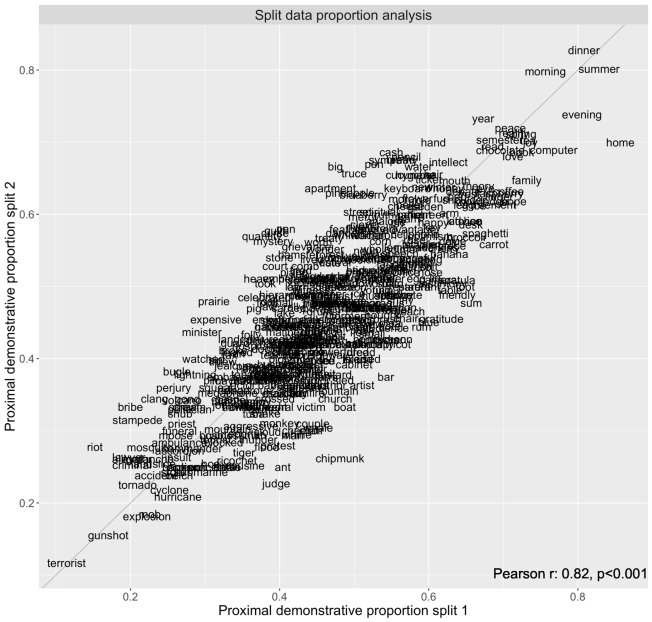

Figure 3. Proportion of proximal demonstratives in two data splits show a high degree of reproducibility (r=0.82) in demonstrative choices

driver of semantic effects in this experiment. This is also clearly visible in figure 4 where semantic features related to manipulability are overshadowed by those related to valence etc. However, manipulability in the previous study (Rocca, et al., 2019a) was defined along three dimensions: "Can you move it with your hands", "Do you want to move it with your hands" and "Will it let you move it with your hands". These dimensions yield a broader definition of manipulability than the one entertained in the present paper, including semantic features that feed into several other factors here, i.e. size, valence/harmfulness and animacy. The manipulability factor used in the present experiment factors out valence and to some degree animacy (into the factor "motion") and the results thus demonstrate that demonstrative choice is affected by manipulability, even with this narrow definition. The broader effect of manipulability is distributed across additional factors that indirectly yield affordances for manipulability.

When combining the effects of the semantic factors and projecting them back into the original feature space, we find that features related to the experiential self dominate (e.g. Needs, Pleasant, Happy) over features related to proximity and the physical self (e.g. Near, Haptic_lan). Whether this effect reflects a hierarchy present outside the experiment or whether it is brought about by the format of the present experimental paradigm remains to be investigated.

This study clearly shows that demonstratives can be used to organize both physical and high-dimensional psychological spaces across an array of semantic dimensions. Taking this line of thought a bit further, we argue that demonstrative choices in the DCT, and perhaps in naturalistic language use, depend on the position of the speaker within the *relevant* feature space, a physical or imaginary hyperspace depending on the context of use.

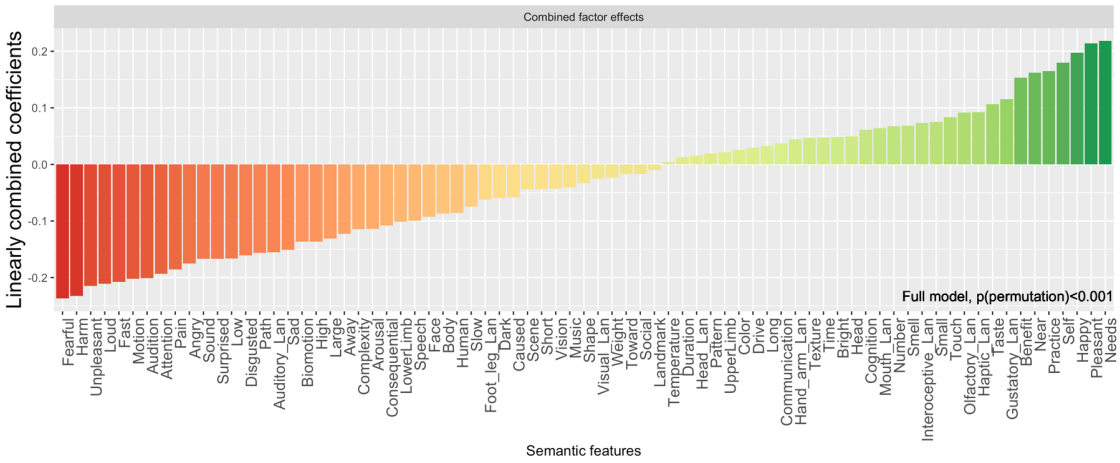

Figure 4. A linear combination of factor loadings and regression coefficients for the 10 significant components shows which semantic features drive of demonstrative choice for proximal (green) and distal (red) demonstratives.

Conclusion

We have found that demonstrative choice is influenced by multiple semantic dimensions, including spatial, bodily and emotional, extending the use of spatial demonstratives beyond physical space to semantic space.

References

Binder, J. R., Conant, L. L., Humphries, C. J., Fernandino, L., Simons, S. B., Aguilar, M., & Desai, R. H. (2016). Toward a brain-based componential semantic representation. *Cognitive Neuropsychology*, 1-45.

Bonfiglioli, C., Finocchiaro, C., Gesierich, B., Rositani, F., & Vescovi, M. (2009). A kinematic approach to the conceptual representations of this and that. *Cognition, 111*, 270-274.

Bühler, K. (1934/2011). *Theory of Language - The representational function of language [Sprachtheorie]*. Amsterdam: John Benjamins.

Caldano, M., & Coventry, K. R. (2019). Spatial demonstratives and perceptual space: To reach or not to reach? *Cognition, 191*, 103989.

Capirci, O., Iverson, J. M., Pizzuto, E., & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of Child Language, 23*, 645-673.

Coventry, K. R., Griffiths, D., & Hamilton, C. J. (2014). Spatial Demonstratives and Perceptual Space: Describing and remembering object location. *Cognitive Psychology*.

Coventry, K. R., Valdés, B., Castillo, A., & Guijarro-Fuentes, P. (2008). Language within your reach: near-far perceptual space and spatial demonstratives. *Cognition, 108*, 889-895.

Diessel, H. (1999). *Demonstratives*: John Benjamins Publishing.

Diessel, H. (2014). Demonstratives, Frames of Reference, and Semantic Universals of Space. *Language and Linguistics Compass, 8*, 116-132.

Hanks, W. F. (2009). Fieldwork on deixis. *Journal of Pragmatics, 41*, 10-24.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*, 179-185.

Leech, G., Rayson, P., & Wilson, A. (2014). *Word Frequencies in Written and Spoken English: based on the British National Corpus*. London: Routledge.

Lynott, D., Connell, L., Brysbaert, M., Brand, J., & Carney, J. (In press). The Lancaster Sensorimotor Norms: Multidimensional measures of Perceptual and Action Strength for 40,000 English words. *Behavior Research Methods*.

Pagel, M., Atkinson, Q. D., S. Calude, A., & Meade, A. (2013). Ultraconserved words point to deep language ancestry across Eurasia. *Proceedings of the National Academy of Sciences, 110*, 8471.

Peeters, D., & Özyürek, A. (2016). This and That Revisited: A Social and Multimodal Approach to Spatial Demonstratives. *Frontiers in Psychology, 7*.

Rocca, R., Coventry, K. R., Tylén, K., Lund, T. E., & Wallentin, M. (in press). Language beyond the language system: dorsal visuospatial pathways support processing of demonstratives and spatial language during naturalistic fast fMRI. *Neuroimage*.

Rocca, R., Tylén, K., & Wallentin, M. (2019a). This shoe, that tiger: Semantic properties reflecting manual affordances of the referent modulate demonstrative use. *PloS One, 14*, e0210333.

Rocca, R., Wallentin, M., Vesper, C., & Tylén, K. (2018). This and that back in context: Grounding demonstrative reference in manual and social affordances. In *Proceedings of The 40th Annual Meeting Of The Cognitive Science Society*. Madison, Wisconsin.

Rocca, R., Wallentin, M., Vesper, C., & Tylén, K. (2019b). This is for you: Social modulations of proximal vs. distal space in collaborative interaction. *Scientific Reports, 9*, 14967.