Kind words

How prosociality shapes cross-modal language

Kees Sommer
Graduation Thesis
Media Technology MSc program,
Leiden University
January 2020
Thesis advisors: Tessa Verhoef and Simon Kirby

Abstract

I propose that self-domestication in humans, correlating with increased activity of neurochemical associated with prosociality, resulted in a heightened sensitivity to cross-modal associations. Evidence supporting this hypothesis is presented. The specific connection between prosociality and cross-modal associations is explored through a behavioural experiment, manipulating participants' prosociality and measuring their sensitivity to cross-modal associations. The resulting data show no significant effect of the manipulation, but do show a significant effect of the employed prosociality measure.

Preface

Those researching language evolution study the earliest origins of human language and protolanguage. These endeavours can be conducted from (a combination of) philosophical, biological, psychological and linguistic perspectives. Charles Darwin is often credited as the first to employ the scientific method on the topic, as he explored it from a biological perspective (Fitch, 2013). The 1990's marked an exponential growth in the field, as computer modeling signaled the way forward (e.g. De Boer, 2000; Kirby, 1999). The field continued to grow and distinct schools emerged, as prominent scholars shifted their focus to experimental research (e.g. Galantucci, 2005; Kirby, Cornish & Smith, 2008; Scott-Phillips, Kirby & Ritchie, 2009).

This research project was firmly embedded within the experimental approach to language evolution. Hypotheses were based on previous findings and real-world observations. The main research question was experimentally investigated in a lab environment with human participants.

This research was motivated by personal experiences with synesthesia, as affected by antidepressant medication. Part of this project consisted of autoethnographic research delving deeper into these events. These experiences are highly subjective and reality-altering and as such can be more successfully reflected on using creative writing techniques (Sommer, 2020).

Acknowledgements

I'd like to thank Tessa Verhoef, Simon Kirby and Christine Cuskley for their supervision of this project. I want to thank Max van Duijn, Teresa Proto and Janelle Ward, who lent me their expert knowledge on social evolution, musical theory and auto-ethnographies. I'd also like to thank Alexander Martin, who kindly recorded the pseudoword stimuli. This project was made possible by an Erasmus+ scholarship provided by the European Commission, through Leiden University and funding from the Creative Intelligence Lab at Leiden University.

1. Introduction

H1: *"Self-domestication in humans, correlating with increased activity of neurochemicals associated with prosociality, resulted in a heightened sensitivity to cross-modal associations. This process helped bootstrap language emergence."*

This hypothesis was based on the following observations:

- Domesticated species display increased activity of neurochemicals associated with prosociality (Calvey, 2019; Herbeck & Gulevich, 2019; Popova, Voitenko, Kulikov & Avgustinovich, 1991);
- Synesthetic experiences can be chemically induced in humans by substances interacting with these neurochemicals (Brang & Ramachandran, 2008; Luke & Terhune, 2013; Terhune, Luke & Kadosh 2017);
- Synesthesia is mediated by the same processes as cross-modal mappings (Bankieris & Simner, 2015; Cuskley, Dingemanse, Kirby & van Leeuwen, 2019; Ramachandran & Hubbard, 2001);
- Cross-modal mappings may have helped bootstrap language emergence (Bankieris & Simner, 2015; Imai & Kita, 2014; Ramachandran & Hubbard, 2001).

I will now explore these observations in more detail.

1.1 Domestication Syndrome

Domestication is the process of altering the genotypic features of a species to favour prosocial behaviour. In most domesticated species (for example, dogs, cats, cows and horses), it is believed that domestication occurred as a result of human influence (Wilkins, Wrangham, & Fitch, 2014), with the possible exception of bonobos which are theorised to be self-domesticated (Hare, Wobber & Wrangham, 2012). In these species, humans would have selectively favoured breeding tamer animals. This process has been experimentally replicated by selectively breeding foxes for tameness, compared to non-selectively breeding them in captivity (Popova et al., 1991).

The term domestication syndrome is used, since there appears to exist a consistent interspecific set of features associated with domestication. Besides increased prosociality, other physiological qualities emerge, such as floppy ears, spotted coats, smaller brains and shorter snouts. Experimentally domesticated subjects also displayed these ancillary features, even though these features had not been selected for (Wilkins et al., 2014). Though the exact mechanisms resulting in these features are not yet fully understood, Wilkins, Wrangham, and Fitch (2014) have related these to neural crest cell deficits. These provide a possible explanation for the emergence of these features.

The reason humans are referred to as (self-)domesticated, is that we also display this set of genotypic features, when compared to our ancestors and closest living genetic relatives. In the case of humans. this process likely resulted from selective pressures of the human social environment. Possibly, the benefits of living in larger groups provided a pressure for prosociality, the equivalent of selecting for tameness (Benítez-Burraco, Theofanopoulou & Boeckx, 2018).

One feature of domestication syndrome is an increase in activity of specific neurochemicals associated with prosociality, such as serotonin and oxytocin (Calvey, 2019; Herbeck & Gulevich, 2019; Popova et al., 1991). Considering the association between these neurochemicals and prosocial behaviour, this result complements the view of domestication as a selection of prosociality (Kiser, Steemer, Branchi & Homberg, 2012). However, it is important to acknowledge that identifying the exact function of any neurochemical is a highly complex task, and labeling one as solely promoting prosociality is too reductionist.

Human self-domestication can already be linked with the evolution of language, as cohabitation in larger groups provides a pressure for communication systems to evolve. Prosocial group interactions, oxytocin and dopamine have further been suggested to facilitate vocal learning (Larsson & Abbott, 2018; Tanaka, Sun, Li & Mooney, 2018; Theofanopoulou, Boeckx & Jarvis, 2017).

1.2 Synesthesia

Synesthesia is a perceptual phenomenon, where certain stimuli elicit sensory experiences on another modality (Ramachandran & Hubbard, 2001). People who consistently experience synesthesia are referred to as congenital synesthetes, or simply synesthetes. The most common form of synesthesia is grapheme-colour synesthesia, where certain graphemes (letters and numbers) are associated with specific colours. However, many other forms are observed. These associations are consistent across extensive periods of time and often operate on constant dimensions (e.g. higher pitches being linearly associated with higher colour frequencies) (Cuskley et al., 2019; Ramachandran & Hubbard, 2001).

Historically, it has been intuited that synesthetes are exposed to “cross-wiring” across areas of the brain associated with the affected modalities. This possibly results from defective synaptic pruning (removal of damaged neurons), which is driven by a genetic mutation (Ramachandran & Hubbard, 2001; Tomson et al., 2011). This mechanism was confirmed by Weiss & Fink (2008) who found increased density of grey matter in the areas of the brain associated with grapheme and colour processing in grapheme-colour synesthetes. Earlier testing in this field had been performed by Esterman, Verstynen, Ivry, and Robertson (2006) who experimentally suppressed synesthetic experiences in synesthetes, by inhibiting sections of their brain.

Besides the research into congenital synesthesia, some preliminary work has been done on synesthesia, as induced through neurochemical manipulation. Many reports provide anecdotal evidence for the induction of synesthesia through chemical substances, such as lysergic acid diethylamide (LSD) and psilocybin (Terhune et al., 2017). One synesthete reported that their synesthetic experiences were exacerbated by the use of bupropion (a norepinephrine-dopamine reuptake inhibitor (NDRI) (Luke & Terhune, 2013)). However, these results are not without their problems. They mostly rely on self-reporting, rather than replicable tests. While many results are anecdotal, these reports are nonetheless consistent across time and between individuals.

Since many of these reports are based on serotonergic interactions, some have suggested serotonin as a specific candidate affecting synesthesia (Brang & Ramachandran, 2008; Luke & Terhune, 2013; Terhune et al., 2017). It is important to realise that the substances reportedly inducing synesthesia interact with a multitude of neurochemicals. Furthemore, oxytocin has been found in relation to other cross-modal phenomena (Zheng et al., 2014). Focusing on only one of these neurochemicals leads to a significant selection bias.

It has been noted that induced synesthesia appears to operate differently from congenital synesthesia. Grapheme-colour synesthesia is relatively underreported in induced synesthesia, compared to the congenital variant. Types of synesthesia reported in chemically induced synesthesia are rarely found in congenital synesthesia (Luke & Terhune, 2013). As such, it is possible that tests designed to identify congenital synesthetes are not appropriate in assessing induced synesthesia. This discrepancy suggests that induced synesthesia results from different mechanisms than congenital synesthesia. However, the experiential similarities between the two, and the fact that neurochemicals can heighten congenital synesthesia suggest that the mechanisms underlying both forms of synesthesia interact.

I would like to propose a mechanical explanation for the interactions between congenital and induced synesthesia. Serotonin, oxytocin and dopamine transmit impulses to the neurons that make up grey matter. An increased activation of these chemicals could also increase the arousal of defectively pruned ("cross-wired") neurons. Since these neurons are more prevalent in synesthetes, they require less activity of these neurochemicals to experience synesthesia (while increasing this activity would increase their synesthetic experiences). If serotonin should indeed be considered a strong contributor to this process, this could relate to the role of serotonin as a potential regulating agent of synaptic plasticity (Kojic, Gu, Douglas & Cynader, 1997). Serotonin could increase the sensitivity of cross-wired neurons to impulses, increasing the likelihood of synesthetic experiences.

1.3 Synesthesia and Cross-modal Mappings

Cross-modal mappings are mappings between sensory modalities. These can present themselves linguistically as cross-modal iconicity, where signals (words) are associated with meanings on other modalities. These words intuitively sound like the meaning they represent. English examples include "tiny", which sounds small, or "zig-zag" which sounds like the movement it refers to.

Cross-modal iconicity has been found to operate consistently across languages. English speakers are able to reliably predict the meaning of japanese ideophones (Iwasaki, Vinson & Vigliocco, 2007) and word-learning accuracy is improved when forms are more iconic (Nielsen & Dingemanse, 2018).

A relationship between synesthesia and cross-modal mappings has previously been intuited (Ramachandran & Hubbard., 2001), since both involve multiple modalities interacting with one-another, and many types of synaesthesia reflect cross-modal patterns intuited by the general population (Bankieris & Simner, 2015). Experimental evidence for this interpretation has been provided by Cuskley, Simner & Kirby (2019) and Bankieris and Simner (2015).

Taking on the "cross-wiring" perspective with regards to synesthesia, it is possible that "cross-wiring" occurs in varying degrees. In this sense, synesthetes would be subject to large amounts of "cross-wired" neurons, while individuals behaving similar to synesthetes would only exhibit some "cross-wiring". This could help explain why some people seem to be more sensitive to chemical induction of synesthesia (Luke & Terhune, 2013).

1.4 Cross-modal Mappings as Bootstrapping

The sound symbolism bootstrapping hypothesis suggests that iconic signals introduced early hominins to the notion that language can refer to something other than itself (Imai & Kita, 2014).

Some suggest that early hominins first imitated real world referents using pantomime, which became increasingly more abstracted (Corballis, 2012). However, this account is less equipped to explain how early hominins encoded referents on other modalities.

The intrinsic association between speech and other modalities (afforded by cross-modal associations) could have introduced the insight that speech sounds can relate to other sensory modalities. By extension, it could have introduced the idea that language can refer to objects and events other than themselves. These shared associations may then have assisted hominins in intuitively building a shared lexicon, that could be intuitively understood by the community (Bankieris & Simner, 2015; Imai & Kita, 2014).

1.5 Summary

Self-domestication in humans, correlating with an increased activity of neurochemicals associated with prosociality (as observed in (experimentally) domesticated species), resulted in increased sensitivity to cross-modal associations (since these chemicals induce synesthesia and synesthesia is mediated by the same processes as cross-modal mappings). This process helped bootstrap language emergence (as cross-modal mappings facilitate cross-modal iconicity and support intuitive meaning sharing).

This hypothesis is still speculative and more research is needed to strengthen its arguments. The biggest rhetorical leap made is the argument that an increased activity of neurochemicals associated with prosociality, resulted in increased sensitivity to cross-modal associations. This is based on the observations that synesthesia seems to be affected by these neurochemicals and that synesthesia is related to cross-modal mappings. Even if these phenomena operate on the same principles, this doesn't mean every single feature is transferable.

With this research project I aimed to provide more insight into this issue, by posing the question:

Q1: *"Does prosociality affect sensitivity to cross-modal associations?"*

This question explores whether the effect of neurochemicals associated with prosociality on synesthesia is transferable to cross-modal mappings. It also shifts the scope from prosocial neurochemicals to prosociality as a behaviour. This was considered a more emergent perspective. Neurochemical research is highly complex and focussing on one neurochemical could be too reductionist.

2. Methods

The design and execution of the experiment took place at the University of Edinburgh. Ethical approval was granted by the University of Edinburgh School of Philosophy, Psychology and Language Sciences. It followed the British Psychological Society Code of Conduct.

2.1 Experimental Design

To help answer the research question, a between participant design was used. Participants were divided into two groups, one test group and one control group. Participants in the test group would perform a pre-task aimed at stimulating prosociality, while participants in the control group would perform a control task. Participants in both groups would then perform a task designed to test their cross-modal sensitivity. Lastly, all participants would fill out a questionnaire measuring their prosocial intentionality, their prosociality and collecting demographic information.

A clapping task was chosen for the pre-task. Synchronising with a rhythm has been shown to increase prosociality in individuals (Kirschner & Tomasello, 2010; von Zimmermann & Richardson, 2016; von von Zimmermann, Vicary, Sperling, Orgs & Richardson, 2018) and these results hold up when reviewed in meta-analysis (Rennung & Göritz, 2016). These tasks have been shown to activate areas of the brain associated with prosociality (Kokal, Engel, Kirschner & Keysers, 2011) and to stimulate the Endogenous Opioid System (Tarr, Launay, Benson & Dunbar, 2017).

The design of the task was based on a methodology by Kokal, Engel, Kirschner & Keysers (2011), where participants were prompted to "drum" a pre-trained rhythm, using an EEG response box. This methodology was chosen since it would be replicable and has shown neuronal activation associated with prosociality. In the original experiment, participants were instructed to "drum" in synchrony with the experimenter in one condition, while participants in a control condition were instructed to "drum" out of synchrony with the experimenter.

This methodology was adapted so that participants in the test condition were clapping along with the experimenter. Participants in the control condition were instructed to clap along with a computer instead.

For the cross-modal task, linguistic stimuli were used to provide a stronger connection between the experiment and the study of language. These were compared with visual features, since there already exists an extensive body of research in this particular association (Cuskley, Simner, & Kirby, 2017; Simner, Cuskley & Kirby, 2010). For the visual features, colour was excluded since this could be complicated by colour blind participants and the multidimensional nature of the data. Roundedness of shape was ignored since this is heavily influenced by the roundedness of the orthographic representation of linguistic stimuli (Cuskley et al., 2017). This resulted in selecting for both visual size and stroke weight.

2.2 Materials

The experiment used custom stimuli in the form of clapped rhythms, computer generated rhythms, pseudowords and the interface used to test participants' cross-modal sensitivity. Besides these custom stimuli, the PBIS questionnaire created by Baumsteiger & Siegel (2019) was used to measure participants' prosocial intentions and prosociality.

2.2.1 Clapped Rhythms

In the test condition, participants would clap along with the experimenter. Participants were asked to clap along to three distinct rhythms. Each rhythm was designed to be easily learnable (since prosociality is stimulated when rhythms come more easily (Kokal et al., 2011)). These rhythms consisted of two measures on a 4/4 meter at a pace of 80 BPM, lasting 6 seconds each. Rhythms were based on capoeira rhythms, since these are intended to be relatively simple to entrain to, while remaining interesting. Multiple rhythms were considered during pilot testing. Two of the resulting rhythms consisted of eight pulses, while the third consisted of eleven pulses.

During testing, the experimenter clapped along to a recording of the rhythms, to maintain consistency between participants. The experimenter would listen to this recording while wearing earbuds. The experimenter would repeat a rhythmic phrase ten times, after which the experimenter paused for twelve seconds (four measures). After each pause, the next rhythm would begin. Blocks consisted of ten repeated phrases of each of the three rhythms, and each block would be repeated three times. This resulted in a task of approximately 10 minutes and 48 seconds (as illustrated in table 1).

	Consisting of	Time
Task	3 blocks	648s.
Block	3 rhythms + 3 pauses (12s. each)	216s.
Rhythm	10 phrases	60s.
Phrase	-	6s.

Table 1: Structure of rhythmic stimuli.

2.2.2 Control Rhythms

In the control condition, participants were instructed to clap along to three distinct rhythms generated by the computer. These rhythms were designed to be difficult to entrain to (since prosociality is more strongly stimulated when synchronising to easier rhythms (kokal et al., 2011)), while keeping duration and amount of pulses consistent with the rhythms in the test condition.

Multiple rhythms were randomly generated using custom software. A set amount of pulses would be randomly placed on a timespan of 6 seconds. A minimum time of 200 milliseconds was ensured between pulses, since this was the shortest amount of time between two pulses in the test condition. Multiple rhythms were generated this way and the ones deemed most difficult to entrain to were selected during piloting.

A separate piece of custom software used the generated timestamps and produced an audible and visual pulse at each interval. Both audible and visual stimuli were presented, since participants in the test condition would be able to see and hear the experimenter clap. The audible pulses were generated from white noise and lasted approximately 175 milliseconds. Noise was preferred over a continuous pitch, since pitch might influence cross-modal associations. The noise was attenuated above 1000 hertz to be more comfortable to listen to. Each pulse was designed to have a short attack and decay time. Visually, the program displayed two black circles close to each other against a grey background (the contrast between the colours was limited to make them more visually comfortable). Prior to each pulse, the two circles would move away from each other, and back to their original position within 200 milliseconds (the time of the shortest possible interval), as illustrated in figure 1 and appendices A and B. This was designed to appear similar to the movement of clapping. The program would repeat each rhythm ten times, before pausing for twelve seconds and generating the next rhythm. The structure of the rhythms was kept consistent with the prosocial condition and as such it is also illustrated in table 1.

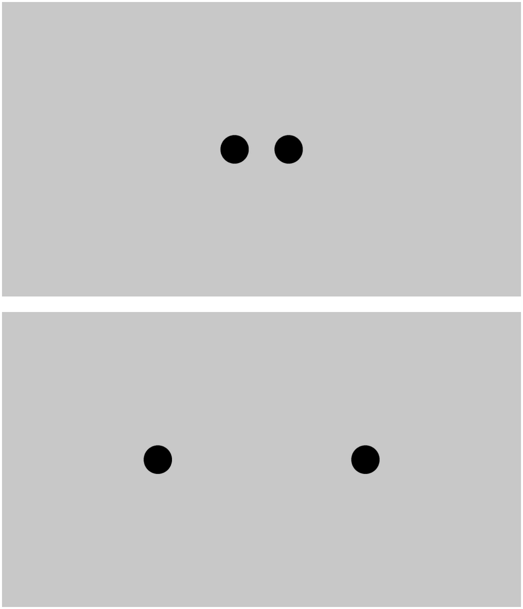

Figure 1: Control condition visual stimuli. The dots move away from each other and back together at each pulse. Top: Dots at their neutral position. Bottom: Dots at their farthest position.

2.2.3 Pseudowords

The linguistic stimuli consisted of pseudowords, to prevent targeting associations with natural language entries. These stimuli were distinctive for both consonant voicing and vowel openness. These features were chosen, since previous findings suggest strong associative reactions between consonant voicing and visual weight, and vowel openness and visual size (Schmidtke, Conrad, & Jacobs, 2014; Simner et al., 2010).

To maintain a pronounced difference in vowel openness, /i/ and /a/ were selected as the tested vowels. To create more variation in the data, two places of articulation were chosen for the consonants (bilabial and alveolar). This resulted in the selection of /p/ /b/ /t/ and /d/ as consonant stimuli. Since it was expected that the strongest association would be found between visual size and vowel openness, a V_1CV_1 structure was used. This resulted in a total of eight distinct pseudowords, as depicted in table 2.

		voiceless	voiced
Closed	bilabial	/ipi/	/ibi/
	alveolar	/iti/	/idi/
Open	bilabial	/apa/	/aba/
	alveolar	/ata/	/ada/

Table 2: Pseudoword stimuli distinguished on three phonological features.

All stimuli were recorded by a trained phonetician (native English speaker) in a dedicated recording studio. A noise filter was applied to the recordings. Since cross-modal associations map more strongly to phonological categories than auditory ones (Cuskley et al., 2019), all pseudowords were recorded to sound natural to a native speaker of British English. Word-initial stress was applied, while keeping stimuli relatively constant for pitch and length. Consonant voicing was partially distinguished by aspiration (of the unvoiced plosives) and word-initial vowels were preceded by a glottal stop.

Since participants would be tested for each pseudoword three times, three recordings were made of each pseudoword. That way participants wouldn't listen to the same recording multiple times. This ensured that participants weren't responding to non-phonological features.

2.2.4 Cross-modal Interface

For the cross-modal task, a custom interface was built, based on code provided by Christine Cuskley. Participants would be presented with a virtual button that prompted a pseudoword to be played. A slider interface would appear simultaneously (as illustrated in figure 2 and appendix C). The slider was shaped as an indented circle, with a semi-rounded shape in its center. Several shapes were considered during piloting and the final design was selected for being neither explicitly sharp or rounded. A circular slider was used, to minimise the interference of left-right biases, while maintaining a one dimensional form of input.

Moving the position of the slider changed both the size and stroke weight of the shape. One end of the slider represented the largest version of the shape, and the other the smallest. The orientation of the slider (the indent on either the right or left side) was randomised between participants, as was the direction that increased the size of the shape (either clockwise or counterclockwise).

Figure 2: Cross-modal interface (right indentation and counterclockwise direction). Left to right: Slider at initial position, at largest position and at smallest position.

2.2.5 Questionnaire

To measure participants' prosocial intentions and prosociality, the PBIS designed by Baumsteiger & Siegel (2019) was used. This questionnaire first presents participants with four prosocial behaviours (such as helping a stranger find something they lost). Participants are asked to rate how likely they are to display each behaviour, on a 7-point Likert scale (1 being very unlikely, 7 being very likely). The mean of their four answers is used as a measure of prosocial intentions. Participants are then asked to type an answer to two open-ended questions. Participants are instructed that these questions are optional and that answering them does not influence their compensation. However, answering them will help the experimenters in their research. The amount of words participants use to answer these questions is interpreted as a measure of prosociality (Baumsteiger & Siegel, 2019).

These tasks were immediately followed by a demographic questionnaire.

2.3 Procedure

The experimenter followed a script to ensure consistency in the procedure. Each participant was invited into the lab and asked to read an information sheet and sign a consent form. Participants received a brief introduction on the type of tasks they would be performing. Each participant was then provided with instructions on the clapping task, differing per condition. Each participant was informed that they would be recorded during the clapping task (for which they had also specifically provided consent). Before the clapping task started, participants were asked if they had any questions.

In the test condition, the experimenter would start the recording of the rhythms on a smartphone, listening to them using ear buds. The experimenter would clap along with the rhythm while sitting opposite of the participant at a distance of approximately one meter. The experimenter would not speak during the clapping task, unless participants had any questions. In the case of the control condition, the experimenter left the lab space during the clapping task, and participants were informed they could ask questions at any time, by knocking on the door.

After participants finished the clapping task, they were provided with instructions on the cross-modal task and questionnaire. The experimenter would leave the lab and participants would complete the experiment by themselves.

During the cross-modal task, participants were presented with three blocks of 8 pseudowords each, meaning participants were presented with each pseudoword three times (24 trials in total). The order of pseudowords was randomly shuffled within each block. After listening to each pseudoword, they were asked to pick the size of the shape that best matched that word, using the cross-modal interface. Instructions to this task remained visible throughout.

At the end of the cross-modal task, participants were directly presented with the questionnaire. Once participants indicated they were done with the experiment, they were debriefed on the goal of the experiment and they were informed about the goal of the open-ended questions. Participants received £5 as compensation and were given the opportunity to ask questions.

3. Results

Sixty-three participants were tested. One participant's data on the cross-modal task and questionnaire were lost due to technical difficulties. Another participant's audio recordings were lost for the same reason. Demographic information was collected on the final 47 participants. Collecting this information required an ethics amendment. Sadly, this was delayed due to a strike at the university.

Of the 47 participants whose demographic data had been collected, the mean age was 23.9 (Min = 18, Max = 31, Med = 23). Ten participants identified as male, thirty-six female and one non-binary. Fourteen participants indicated having studied linguistics, or a closely related field. Eight participants indicated that they had previously experienced synesthesia, versus Nineteen who indicated they hadn't and Twenty who were uncertain. Of the 62 participants whose cross-modal and prosocial data was usable, 31 participated in either condition. The mirroring of the cross-modal interface was applied to half of participants. Thirty participants were in the clockwise condition, versus thirty-two in the counterclockwise condition.

Since size and weight were reflected using the same measure, both were referred to as the size variable. Vowel openness was coded as + or - closed, place of articulation was coded as + or - labial and voicing was coded as + or - voiced.

I tested whether phonological features predicted size and stroke weight, in order to test if these expected preferences were significant in the data. A linear mixed effects model was fit to the data, where size was predicted by these phonological features and participant ID was used as a random variable to account for personal preferences and mirroring conditions (size ~ closed + voiced + labial + (1|ID)).

This model was compared to three reduced versions of the model, with one of the phonological features omitted in each of the models, using anova testing. This showed that each phonological feature had a significant effect on the model (openness: $p < .001$; place of articulation: $p = .01$; voicing: $p = .02$). Figure 3 illustrates the strength of each phonological feature, as the sizes assigned to each word displaying that feature are plotted.

Having shown that participants significantly associated the phonological features with size and weight, I tested the extent to which this was influenced by participant condition. A ranking was assigned to each word, to indicate their expected size, based on literature. Participants were expected to map closed vowels (/i/) and voiceless consonants (/t/ /p/) with smaller sizes and stroke weights, versus open vowels (/a/) and voiced consonants (/d/ /b/) with larger sizes and weights (Schmidtke et al., 2014; Simner et al., 2010). No strong relationship between size or weight and place of articulation was expected. Vowel openness was expected to have a stronger effect than consonant voicing. This resulted in the following expected size rating for each word (ranging from 0 to 1): /aba/, /ada/ = 1; /apa/, /ata/ = 0.67; /ibi/, /idi/ = 0.33; /ipi/ /iti/ = 0.

A linear mixed-effects model was fit to predict size, based on the expected size rating and participant condition, while once

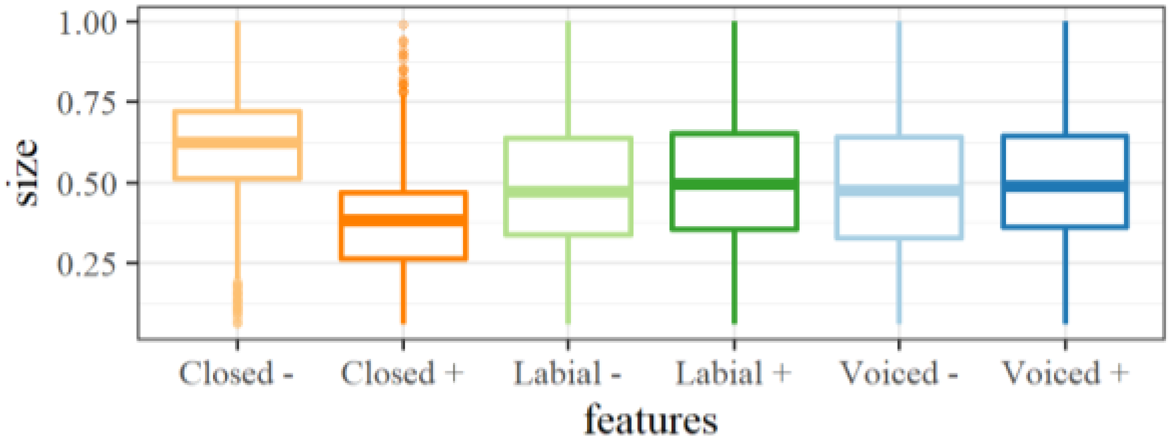

Figure 3: Sizes assigned to pseudowords displaying each feature

again using participant ID as a random variable (size ~ condition * expected size + (1|ID)). This model was compared to a reduced version, where the interaction between the condition and the expected size was omitted, using anova testing. This showed that there was no significant effect of condition on size (p = .14).

The next step was to investigate how size was affected by the prosociality measure. To better understand the data, it was visualised in figure 4. Here, the size picked for each pseudoword was plotted against the prosociality score of the participant who picked it. A linear model was calculated for the sizes assigned to each pseudoword, as depicted by the lines. Intuitively, it appears that, as prosociality increases, the sizes assigned to the pseudowords become more distinctive. This is most strongly illustrated by lines representing /a/ words moving up, and those representing /i/ words moving down, as

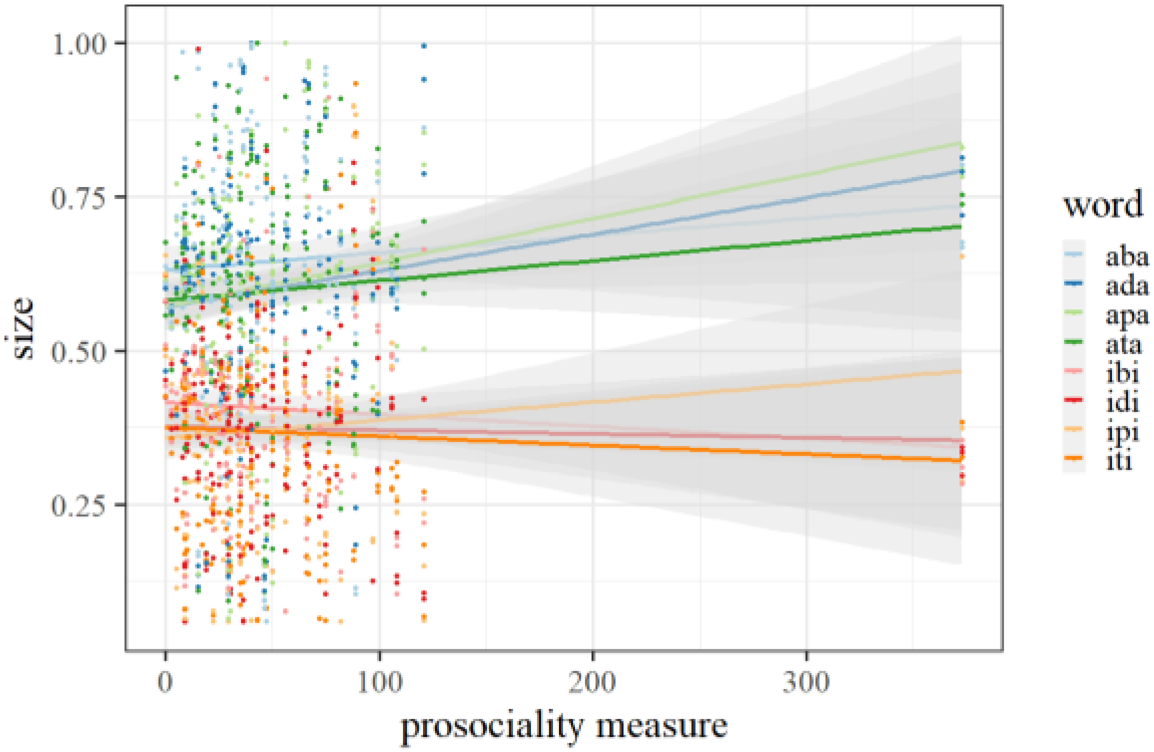

Figure 4: Size/weight for prosociality per pseudoword

prosociality increases. However, it also illustrates the importance of controlling for participants, as there exists a large gap between the highest and second highest prosociality measures.

A linear mixed-effects model was again fit to the data, based on expected size rating and participant's prosociality measure (word count), while controlling for prosociality per participant as a random variable (size ~ prosociality * expected size + (1|ID) + (prosociality-1|ID)). The random slope and intercept were not expected to correlate. This model was compared to a reduced version where the interaction between prosociality and expected size was omitted, using anova testing. This showed a significant difference (p = .03), meaning that the interaction between these variables had a significant effect on the model.

4. Discussion

The results of this experiment show a number of interesting things. Firstly, it would appear that the pre-task manipulation (clapping with the experimenter vs. clapping along with the computer), did not generate a measurable difference in performance on the cross-modal task.

It is of course very possible that neither pre-task had an effect on prosociality. It is also possible that the conditions were too similar to generate a measurable difference, and influenced prosociality in comparable ways. It is possible that the difference in difficulty wasn't strong enough, and that the presence of another person synchronising with participants didn't make a measurable difference.

However, the prosociality measure did have a significant effect on selected size. This finding supports the hypothesis that prosociality positively influences cross-modal sensitivity.

The fact that the pre-task didn't generate a measurable difference between participants, while the prosociality measure did presents an interesting topic for future research. Speculatively, this could be explained by the fact that some neurochemicals behave more stable than others. For example, while mu-opioids are more sensitive to prosocial manipulations (Manninen et al., 2017), serotonin is much more stable. Serotonin reuptake inhibitor therapy has a delayed effect ranging from several days to multiple weeks (Mitchell, 2006). Since the prosocial manipulation did not measurably influence cross-modal sensitivity but prosociality did, the effect could more strongly relate to more stable neurochemical interactions. In this report I have previously highlighted the potential for serotonin in particular as an agent affecting synesthesia.

Currently, there are some complications with the analysis. Prosociality (as in the original methodology) was interpreted as an unbound variable. This resulted in a large difference between the highest and second highest prosociality measure. Future research should consider whether an unbound variable is appropriate when measuring prosociality.

Furthermore, phonological features were interpreted as non-interacting. Though the phonetic realisations of these features interact, this project focused on their phonological representation. However, alternative interpretations are possible.

In the adopted model, vowel openness was expected to have a larger effect than consonant voicing, but the increments between expected sizes were kept consistent. This was perceived as an agnostic approach, since minimal assumptions on the strength of each effect are made, but alternative measures can be imagined.

Word-count was used as a test of prosociality, since this was the method employed in the original model by Baumsteiger & Siegel (2019). They already found this measure interacted with prosocial intentions, and therefore identified it as an appropriate measure of prosociality. However, taken in isolation, the measure is impossible to disconnect from eloquence, "wordiness" or time available to the participant. All these things could have their own interactions with prosociality or performance on the cross-modal task. Behavioural tests for prosociality generally pose methodological challenges. Prosociality is a broad range of complex behaviours and as such difficult to quantify. Though I believe the selected test to be sufficient for a preliminary research, it should be improved upon in further research.

The difficulty of measuring prosociality provides more motivation to investigate specific neurochemicals. Though these present their own methodological challenges, they could help avoid the vagueness of measuring prosociality.

Lastly, though the topics researched interact with medicinal use, recreational drug usage and mental health issues, no demographic data was collected on these topics. Future research on the topic would benefit from taking these aspects into consideration.

5. Conclusion

This research project introduced the following hypothesis:

"Self-domestication in humans, correlating with increased prosocial neurochemical activity, resulted in a heightened sensitivity to cross-modal associations. This process helped bootstrap language emergence."

The research question "Does prosociality affect sensitivity to cross-modal associations?" was studied. Prosociality was approached as a behavioural phenomenon, rather than a neurochemical one. I addressed this question using a behavioural experiment, attempting to manipulate participants' prosociality through a clapping task and measuring their cross-modal sensitivity and prosociality. The results of this experiment supported the notion that prosociality affects cross-modal sensitivity. These findings build upon the proposed hypothesis, by connecting prosociality with cross-modal associations. This connection had been inferred from the literature, but required further investigation.

The goal of this project was to test if the effect of neurochemicals associated with prosociality on synesthesia could be observed in cross-modal sensitivity, using behavioural methods. However, supporting the notion that the origins of this interaction are neurochemical requires the manipulation or measurement of specific neurochemicals.

A proposed follow-up study could focus on more specific neurochemicals, such as serotonin. One approach could involve recruiting participants currently treated with selective serotonin reuptake inhibitors (SSRIs). Their sensitivity to cross-modal associations could be measured and compared to participants from a control population. More complex interactions between neurochemicals could be studied by working with participants under the influence of the chemical agents that reportedly induce synesthesia (such as psilocybin). Comparing their results on cross-modal tests to those of controls could help inform the role of neurochemicals in complex interactions.

Overall, this project has merely scratched the surface on prosociality, not just as a pressure on the evolution of language, but as a neurological mechanism facilitating intuitive meaning sharing.

References

Bankieris, K., & Simner, J. (2015). What is the link between synaesthesia and sound symbolism? *Cognition, 136*, 186-195.

Baumsteiger, R., & Siegel, J. T. (2019). Measuring prosociality: The development of a prosocial behavioral intentions scale. *Journal of personality assessment*, *101*(3), 305-314.

Benítez-Burraco, A., Theofanopoulou, C., & Boeckx, C. (2018). Globularization and domestication. *Topoi*, *37*(2), 265-278.

De Boer, B. (2000). Self-organization in vowel systems. *Journal of phonetics*, *28*(4), 441-465.

Brang, D., & Ramachandran, V. S. (2008). Psychopharmacology of synesthesia; the role of serotonin S2a receptor activation. *Medical Hypotheses*, *70*(4), 903.

Calvey, T. (2019). Human Self-Domestication and the Extended Evolutionary Synthesis of Addiction: How Humans Evolved a Unique Vulnerability. *Neuroscience, 419*, 100-107.

Corballis, M. C. (2012). How language evolved from manual gestures. *Gesture*, *12*(2), 200-226.

Cuskley, C., Dingemanse, M., Kirby, S., & van Leeuwen, T. M. (2019). Cross-modal associations and synesthesia: Categorical perception and structure in vowel–color mappings in a large online sample. *Behavior research methods*, 1-25.

Cuskley, C., Simner, J., & Kirby, S. (2017). Phonological and orthographic influences in the bouba–kiki effect. *Psychological research*, *81*(1), 119-130.

Esterman, M., Verstynen, T., Ivry, R. B., & Robertson, L. C. (2006). Coming unbound: disrupting automatic integration of synesthetic color and graphemes by transcranial magnetic stimulation of the right parietal lobe. *Journal of cognitive neuroscience*, *18*(9), 1570-1576.

Fitch, W. T. (2013). Musical protolanguage: Darwin's theory of language evolution revisited. *Birdsong, speech, and language: Exploring the evolution of mind and brain*, 489-503.

Galantucci, B. (2005). An experimental study of the emergence of human communication systems. *Cognitive science*, *29*(5), 737-767.

Hare, B., Wobber, V., & Wrangham, R. (2012). The self-domestication hypothesis: evolution of bonobo psychology is due to selection against aggression. *Animal Behaviour*, *83*(3), 573-585.

Herbeck, Y., & Gulevich, E. (2019). Neuropeptides as facilitators of domestication. *Cell and Tissue Research, 375*(1), 295-307.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 369*(1651), 20130298.

Iwasaki, N., Vinson, D. P. & Vigliocco, G (2007). What do English Speakers Know about gera-gera and yota-yota?: A Cross-linguistic Investigation of Mimetic Words of Laughing and Walking. *Japanese language education around the globe; Japanese language education around the globe*, *17*, 53-78.

Kirby, S. (1999). Syntax out of learning: The cultural evolution of structured communication in a population of induction algorithms. In

European Conference on Artificial Life (pp. 694-703). Springer, Berlin, Heidelberg.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681-10686.

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354-364.

Kiser, D., SteemerS, B., Branchi, I., & Homberg, J. R. (2012). The reciprocal interaction between serotonin and social behaviour. *Neuroscience & Biobehavioral Reviews*, *36*(2), 786-798.

Kojic, L., Gu, Q., Douglas, R. M., & Cynader, M. S. (1997). Serotonin facilitates synaptic plasticity in kitten visual cortex: an in vitro study. *Developmental Brain Research*, *101*(1-2), 299-304.

Kokal, I., Engel, A., Kirschner, S., & Keysers, C. (2011). Synchronized drumming enhances activity in the caudate and facilitates prosocial commitment-if the rhythm comes easily. *PLoS One*, *6*(11), e27272.

Larsson, M., & Abbott, B. (2018). Is the Capacity for Vocal Learning in Vertebrates Rooted in Fish Schooling Behavior? *Evolutionary Biology, 45*(4), 359-373.

Luke, D., & Terhune, D. B. (2013). The induction of synaesthesia with chemical agents: a systematic review. *Frontiers in psychology*, *4*, 753.

Manninen, S., Tuominen, L., Dunbar, R. I., Karjalainen, T., Hirvonen, J., Arponen, E. & Nummenmaa, L. (2017). Social laughter triggers endogenous opioid release in humans. *Journal of Neuroscience*, *37*(25), 6125-6131.

Mitchell, A. J. (2006). Two-week delay in onset of action of antidepressants: new evidence. *The British Journal of Psychiatry*, *188*(2), 105-106.

Nielsen, A., & Dingemanse, M. (2018). Iconicity in word learning and beyond: A critical review.

Popova, N. K., Voitenko, N. N., Kulikov, A. V., & Avgustinovich, D. F. (1991). Evidence for the involvement of central serotonin in mechanism of domestication of silver foxes. *Pharmacology Biochemistry and Behavior*, *40*(4), 751-756.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia--a window into perception, thought and language. *Journal of consciousness studies*, *8*(12), 3-34.

Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony. *Zeitschrift für Psychologie*.

Schmidtke, D., Conrad, M., & Jacobs, A. M. (2014). Phonological iconicity. *Frontiers in psychology*, *5*, 80.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*(2), 226-233.

Simner, J., Cuskley, C., & Kirby, S. (2010). What sound does that taste? Cross-modal mappings across gustation and audition. *Perception*, *39*(4), 553-569.

Sommer, K. (2020). A strange waking. Unpublished manuscript.

Tanaka, M., Sun, F., Li, Y., & Mooney, R. (2018). A mesocortical dopamine circuit enables the cultural transmission of vocal behaviour. *Nature, 563*(7729), 117-120.

Tarr, B., Launay, J., Benson, C., & Dunbar, R. I. (2017). Naltrexone blocks endorphins released when dancing in synchrony. *Adaptive Human Behavior and Physiology*, *3*(3), 241-254.

Terhune, D. B., Luke, D. P., & Kadosh, R. C. (2017). The induction of synaesthesia in non-synaesthetes. *Sensory Blending: On Synaesthesia and Related Phenomena*, 215.

Theofanopoulou, C., Boeckx, C., & Jarvis, E. (2017). A hypothesis on a role of oxytocin in the social mechanisms of speech and vocal learning. *Proceedings. Biological Sciences, 284*(1861), 20170988-20170988.

Tomson, S. N., Avidan, N., Lee, K., Sarma, A. K., Tushe, R., Milewicz, D. M., ... & Eagleman, D. M. (2011). The genetics of colored sequence synesthesia: suggestive evidence of linkage to 16q and genetic heterogeneity for the condition. *Behavioural brain research*, *223*(1), 48-52.

Weiss, P. H., & Fink, G. R. (2008). Grapheme-colour synaesthetes show increased grey matter volumes of parietal and fusiform cortex. *Brain*, *132*(1), 65-70.

Wilkins, A. S., Wrangham, R. W., & Fitch, W. T. (2014). The “domestication syndrome” in mammals: a unified explanation based on neural crest cell behavior and genetics. *Genetics*, *197*(3), 795-808.

Zheng, J. J., Li, S. J., Zhang, X. D., Miao, W. Y., Zhang, D., Yao, H., & Yu, X. (2014). Oxytocin mediates early experience–dependent cross-modal plasticity in the sensory cortices. *Nature neuroscience*, *17*(3), 391.

von Zimmermann, J., & Richardson, D. C. (2016). Verbal synchrony and action dynamics in large groups. *Frontiers in Psychology*, *7*, 2034.

von Zimmermann, J., Vicary, S., Sperling, M., Orgs, G., & Richardson, D. C. (2018). The choreography of group affiliation. *Topics in Cognitive Science*, *10*(1), 80-94.

Appendices

A: Control task stimuli

https://drive.google.com/open?id=196NONA7y0UikpfqJPWqbZuDlqosD-A4-

B: Control task stimuli demonstration

https://drive.google.com/open?id=1SSs5_SEWG3xLSbKeGz4l0W-jT5dFa7jg

C: Cross-modal interface demonstration

https://drive.google.com/open?id=127iQUVWoYxVBF6r31YDDzVQNKxgEz1Xu